(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
6 February 2003 (06.02.2003)

PCT

(10) International Publication Number
WO 03/010194 A2

(51) International Patent Classification[7]: C07K 14/195

(21) International Application Number: PCT/IB02/03396

(22) International Filing Date: 26 July 2002 (26.07.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

0118401.9	27 July 2001 (27.07.2001)	GB
0121591.2	6 September 2001 (06.09.2001)	GB
0211025.2	14 May 2002 (14.05.2002)	GB

(71) Applicant *(for all designated States except US)*: CHIRON SPA [IT/IT]; Via Fiorentina, 1, I-53100 Siena (IT).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: ARICO, Maria [IT/IT]; Chiron SpA, Via Fiorentina, 1, I-53100 Siena (IT). COMANDUCCI, Maurizio [IT/IT]; Chiron SpA, Via Fiorentina, 1, I-53100 Siena (IT).

(74) Agents: MARSHALL, Cameron, John et al.; Carpmaels & Ransford, 43 Bloomsbury Square, London WC1A 2RA (GB).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

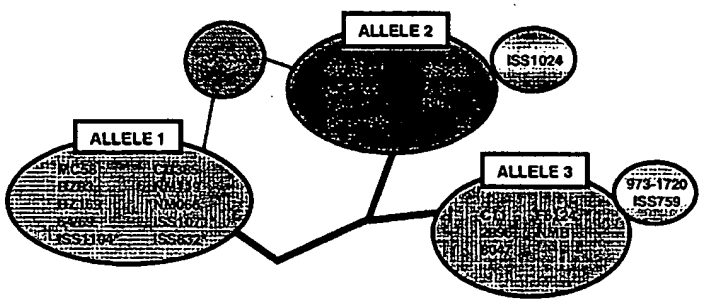
(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

[Continued on next page]

(54) Title: MENINGOCOCCUS ADHESINS



(57) Abstract: NadA, App and ORF40 function as adhesins in *N.meningitidis*. Adhesion can be modulated by targeting these three proteins. NadA allelic variants are disclosed. Autoproteolytic cleavage of App is disclosed, as is removal of the activity by mutagenesis. App is processed and secreted into culture medium when expressed in *E.coli*. Mature App proteins are disclosed. Knockout mutants are disclosed. Vesicles from non-Neisserial hosts with heterologous adhesin expression are disclosed.

WO 03/010194 A2

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

MENINGOCOCCUS ADHESINS

All documents cited herein are incorporated by reference in their entirety.

TECHNICAL FIELD

This invention is in the field of biochemistry and, in particular, the biochemistry of the pathogenic bacteria in the genus *Neisseria* (*e.g. N.meningitidis* and *N.gonorrhoeae*).

BACKGROUND ART

International patent applications WO99/24578, WO99/36544, WO99/57280 and WO00/22430 disclose proteins from *Neisseria meningitidis* and *Neisseria gonorrhoeae*. The complete genome sequence of serogroup B *N.meningitidis* has been published [Tettelin *et al.* (2000) *Science* 287:1809-1815] and has been subjected to analysis in order to identify vaccine antigens [Pizza *et al.* (2000) *Science* 287:1816-1820]. Approaches to expression of the proteins are disclosed in WO01/64922. The complete genome sequence of serogroup A *N.meningitidis* is also known [Parkhill *et al.* (2000) *Nature* 404:502-506].

Sequence data alone, however, does not reveal everything about this pathogen. Objects of the present invention include: (a) to provide ways of intervening in *Neisseria* biochemistry; (b) to provide new uses for known *Neisseria* proteins; (c) to provide alternative and improved forms of known *Neisseria* proteins, such as enzymatically inactive forms of known proteins or proteolytic products of known proteins; and (d) to provide materials useful for studying and modulating Neisserial adhesion.

DISCLOSURE OF THE INVENTION

Nomenclature used herein

'ORF40' is disclosed in example 1 of WO99/36544. Sequences from serogroups A and B of *N.meningitidis* are disclosed (SEQ IDs 1 to 6 therein). Other forms of the protein are disclosed in WO99/31132 and WO99/58683, and can also be found in GenBank (see gi accession numbers: 11352902, 7228562, 14578015, 12958107, 7228586, 7228572, 7228594, 7228588, 14578013, 7228568, 7228546, 7228548, 7228592, 14578009, 7228558, 7228600, 7228596, 7228542, 7228574, 7228552, 7228554, 14578023, 14578021, 11354080, 7228584 & 7228590).

'App' (adhesion and penetration protein) is disclosed as 'ORF1' in example 77 of WO99/24578. Sequences from serogroups A and B of *N.meningitidis* and from *N.gonorrhoeae* are disclosed (SEQ IDs 647 to 654 therein). Other forms of the protein are disclosed in WO99/55873, and can also be found in GenBank (see gi accession numbers: 11280386, 7227246, 11071865, 6977941, 11071863, 11280387, 7379205).

'NadA' (Neisserial adhesin A) from serogroup B of *N.meningitidis* is disclosed as protein '961' in WO99/57280 (SEQ IDs 2943 & 2944) and as 'NMB1994' by Tettelin *et al.* (see also GenBank accession numbers: 11352904 & 7227256) and in Figure 9 herein.

These proteins are preferably expressed other than as a fusion protein (*e.g.* without GST, MBP, his-tag or similar).

Preferred proteins for use according to the invention are those of serogroup B *N.meningitidis* strain MC58, strain 2996 or strain 394/98 (a New Zealand strain). It will be appreciated, however, that the invention is not in general limited by strain – references to a particular protein (*e.g.* 'ORF40', 'App' *etc.*) may be taken to include that protein from any strain. In general, therefore, reference to any particular protein includes proteins which share sequence identity with one of the sequences disclosed above. The degree of 'sequence identity' is preferably greater than 50% (*eg.* 60%, 70%, 80%, 90%, 95%, 99% or more). This includes mutants and allelic variants. In the context of the present invention, sequence identity is preferably determined by the Smith-Waterman homology search algorithm as implemented in the MPSRCH program (Oxford Molecular), using an affine gap search with parameters *gap open penalty=12* and *gap extension penalty=1*. Typically, 50% identity or more between two proteins is considered to be an indication of functional equivalence.

The naming conventions used in WO99/24578, WO99/36544 and WO99/57280 are also used herein (*e.g.* 'ORF4', 'ORF40', 'ORF40-1' *etc.* as used in WO99/24578 and WO99/36544; 'm919', 'g919' and 'a919' *etc.* as used in WO99/57280).

Secreted App

It has been found that, when expressed in *E.coli* without a GST or his-tag fusion partner, App is exported to the outer membrane as a precursor of about 160kDa, where it is processed and secreted into the culture.

The invention therefore provides a method for purifying processed App protein, comprising the steps of: expressing a gene encoding App protein in a non-Neisserial host cell; and purifying processed App protein from the culture medium.

The invention also provides purified protein obtainable by this process.

The App protein preferably includes its wild-type 42 residue signal peptide at the N-terminus *i.e.* no N-terminus fusion partner is used. It is also preferred not to include a C-terminus fusion partner.

To purify the protein from the culture medium, the culture can be centrifuged and the protein can be recovered from the supernatant.

The non-Neisserial host cell is preferably a bacterium and is most preferably *E.coli*.

Bacterial expression techniques are known in the art. A bacterial promoter is any DNA sequence capable of binding bacterial RNA polymerase and initiating the downstream (3') transcription of a coding sequence (*eg.* structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site and a transcription initiation site. A bacterial promoter may also have a second domain called an operator, that may overlap an adjacent RNA polymerase binding site at which RNA synthesis begins. The operator permits negative regulated (inducible) transcription, as a gene repressor protein may bind the operator and thereby inhibit transcription of a specific gene. Constitutive expression may occur in the absence of negative regulatory elements, such as the operator. In addition, positive regulation may be achieved by a gene

activator protein binding sequence, which, if present is usually proximal (5') to the RNA polymerase binding sequence. An example of a gene activator protein is the catabolite activator protein (CAP), which helps initiate transcription of the lac operon in Escherichia coli (E. coli) [Raibaud *et al.* (1984) *Annu. Rev. Genet. 18*:173]. Regulated expression may therefore be either positive or negative, thereby either enhancing or reducing transcription.

Sequences encoding metabolic pathway enzymes provide particularly useful promoter sequences. Examples include promoter sequences derived from sugar metabolizing enzymes, such as galactose, lactose (*lac*) [Chang *et al.* (1977) *Nature 198*:1056], and maltose. Additional examples include promoter sequences derived from biosynthetic enzymes such as tryptophan (*trp*) [Goeddel *et al.* (1980) *Nuc. Acids Res. 8*:4057; Yelverton *et al.* (1981) *Nucl. Acids Res. 9*:731; US patent 4,738,921; EP-A-0036776 and EP-A-0121775]. The g-laotamase (*bla*) promoter system [Weissmann (1981) "The cloning of interferon and other mistakes." In *Interferon 3* (ed. I. Gresser)], bacteriophage lambda PL [Shimatake *et al.* (1981) *Nature 292*:128] and T5 [US patent 4,689,406] promoter systems also provide useful promoter sequences.

In addition, synthetic promoters which do not occur in nature also function as bacterial promoters. For example, transcription activation sequences of one bacterial or bacteriophage promoter may be joined with the operon sequences of another bacterial or bacteriophage promoter, creating a synthetic hybrid promoter [US patent 4,551,433]. For example, the *tac* promoter is a hybrid *trp-lac* promoter comprised of both *trp* promoter and *lac* operon sequences that is regulated by the *lac* repressor [Amann *et al.* (1983) *Gene 25*:167; de Boer *et al.* (1983) *Proc. Natl. Acad. Sci. 80*:21]. Furthermore, a bacterial promoter can include naturally occurring promoters of non-bacterial origin that have the ability to bind bacterial RNA polymerase and initiate transcription. A naturally occurring promoter of non-bacterial origin can also be coupled with a compatible RNA polymerase to produce high levels of expression of some genes in prokaryotes. The bacteriophage T7 RNA polymerase/promoter system is an example of a coupled promoter system [Studier *et al.* (1986) *J. Mol. Biol. 189*:113; Tabor *et al.* (1985) *Proc Natl. Acad. Sci. 82*:1074]. In addition, a hybrid promoter can also be comprised of a bacteriophage promoter and an *E. coli* operator region (EPO-A-0 267 851).

In addition to a functioning promoter sequence, an efficient ribosome binding site is also useful for the expression of foreign genes in prokaryotes. In *E. coli*, the ribosome binding site is called the Shine-Dalgarno (SD) sequence and includes an initiation codon (ATG) and a sequence 3-9 nucleotides in length located 3-11 nucleotides upstream of the initiation codon [Shine *et al.* (1975) *Nature 254*:34]. The SD sequence is thought to promote binding of mRNA to the ribosome by the pairing of bases between the SD sequence and the 3' and of *E. coli* 16S rRNA [Steitz *et al.* (1979) "Genetic signals and nucleotide sequences in messenger RNA." In *Biological Regulation and Development: Gene Expression* (ed. R.F. Goldberger)]. To express eukaryotic genes and prokaryotic genes with weak ribosome-binding site [Sambrook *et al.* (1989) "Expression of cloned genes in Escherichia coli." In *Molecular Cloning: A Laboratory Manual*].

A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus will always be a methionine, which is encoded by the ATG start codon. If

desired, methionine at the N-terminus may be cleaved from the protein by *in vitro* incubation with cyanogen bromide or by either *in vivo* on *in vitro* incubation with a bacterial methionine N-terminal peptidase (EP-A-0219237).

Usually, transcription termination sequences recognized by bacteria are regulatory regions located 3' to the translation stop codon, and thus together with the promoter flank the coding sequence. These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA. Transcription termination sequences frequently include DNA sequences of about 50 nucleotides capable of forming stem loop structures that aid in terminating transcription. Examples include transcription termination sequences derived from genes with strong promoters, such as the *trp* gene in *E. coli* as well as other biosynthetic genes.

Usually, the above described components, comprising a promoter, signal sequence (if desired), coding sequence of interest, and transcription termination sequence, are put together into expression constructs. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (*eg.* plasmids) capable of stable maintenance in a host, such as bacteria. The replicon will have a replication system, thus allowing it to be maintained in a prokaryotic host either for expression or for cloning and amplification. In addition, a replicon may be either a high or low copy number plasmid. A high copy number plasmid will generally have a copy number ranging from about 5 to about 200, and usually about 10 to about 150. A host containing a high copy number plasmid will preferably contain at least about 10, and more preferably at least about 20 plasmids. Either a high or low copy number vector may be selected, depending upon the effect of the vector and the foreign protein on the host.

Alternatively, the expression constructs can be integrated into the bacterial genome with an integrating vector. Integrating vectors usually contain at least one sequence homologous to the bacterial chromosome that allows the vector to integrate. Integrations appear to result from recombinations between homologous DNA in the vector and the bacterial chromosome. For example, integrating vectors constructed with DNA from various Bacillus strains integrate into the Bacillus chromosome (EP-A-0127328). Integrating vectors may also be comprised of bacteriophage or transposon sequences.

Usually, extrachromosomal and integrating expression constructs may contain selectable markers to allow for the selection of bacterial strains that have been transformed. Selectable markers can be expressed in the bacterial host and may include genes which render bacteria resistant to drugs such as ampicillin, chloramphenicol, erythromycin, kanamycin (neomycin), and tetracycline [Davies *et al.* (1978) *Annu. Rev. Microbiol. 32*:469]. Selectable markers may also include biosynthetic genes, such as those in the histidine, tryptophan, and leucine biosynthetic pathways.

Alternatively, some of the above described components can be put together in transformation vectors. Transformation vectors are usually comprised of a selectable market that is either maintained in a replicon or developed into an integrating vector, as described above.

Expression and transformation vectors, either extra-chromosomal replicons or integrating vectors, have been developed for transformation into many bacteria. For example, expression vectors have been developed for, *inter alia*, the following bacteria: Bacillus subtilis [Palva *et al.* (1982) *Proc. Natl. Acad. Sci. USA 79*:5582; EP-A-0 036 259 and EP-A-0 063 953; WO 84/04541], Escherichia coli [Shimatake *et al.* (1981) *Nature 292*:128; Amann *et al.* (1985) *Gene 40*:183; Studier *et al.* (1986) *J. Mol. Biol. 189*:113; EP-A-0 036 776,EP-A-0 136 829 and EP-A-0 136 907], Streptococcus cremoris [Powell *et al.* (1988) *Appl. Environ. Microbiol. 54*:655]; Streptococcus lividans [Powell *et al.* (1988) *Appl. Environ. Microbiol. 54*:655], Streptomyces lividans [US patent 4,745,056].

Methods of introducing exogenous DNA into bacterial hosts are well-known in the art, and usually include either the transformation of bacteria treated with $CaCl_2$ or other agents, such as divalent cations and DMSO. DNA can also be introduced into bacterial cells by electroporation. Transformation procedures usually vary with the bacterial species to be transformed. See *eg.* [Masson *et al.* (1989) *FEMS Microbiol. Lett. 60*:273; Palva *et al.* (1982) *Proc. Natl. Acad. Sci. USA 79*:5582; EP-A-0 036 259 and EP-A-0 063 953; WO 84/04541, Bacillus], [Miller *et al.* (1988) *Proc. Natl. Acad. Sci. 85*:856; Wang *et al.* (1990) *J. Bacteriol. 172*:949, Campylobacter], [Cohen *et al.* (1973) *Proc. Natl. Acad. Sci. 69*:2110; Dower *et al.* (1988) *Nucleic Acids Res. 16*:6127; Kushner (1978) "An improved method for transformation of Escherichia coli with ColE1-derived plasmids. In *Genetic Engineering: Proceedings of the International Symposium on Genetic Engineering* (eds. H.W. Boyer and S. Nicosia); Mandel *et al.* (1970) *J. Mol. Biol. 53*:159; Taketo (1988) *Biochim. Biophys. Acta 949*:318; Escherichia], [Chassy *et al.* (1987) *FEMS Microbiol. Lett. 44*:173 Lactobacillus]; [Fiedler *et al.* (1988) *Anal. Biochem 170*:38, Pseudomonas]; [Augustin *et al.* (1990) *FEMS Microbiol. Lett. 66*:203, Staphylococcus], [Barany *et al.* (1980) *J. Bacteriol. 144*:698; Harlander (1987) "Transformation of Streptococcus lactis by electroporation, in: *Streptococcal Genetics* (ed. J. Ferretti and R. Curtiss III); Perry *et al.* (1981) *Infect. Immun. 32*:1295; Powell *et al.* (1988) *Appl. Environ. Microbiol. 54*:655; Somkuti *et al.* (1987) *Proc. 4th Evr. Cong. Biotechnology 1*:412, Streptococcus].

Adherence proteins

Example 22 of international patent application WO01/64922 discloses that *E.coli* which expresses protein NadA can adhere to human epithelial cells. This adherence activity has been further studied and it has also been found for App and ORF40.

The invention provides methods for preventing the attachment of Neisserial cells to epithelial cells.

References to a "Neisserial cell" in this section include any species of the bacterial genus *Neisseria*, including *N.gonorrhoeae* and *N.lactamica*. Preferably, however, the species is *N.meningitidis*. The *N.meningitidis* may be from any serogroup, including serogroups A, C, W135 and Y. Most preferably, however, it is *N.meningitidis* serogroup B.

References to an "epithelial cell" in this section include any cell found in or derived from the epithelium of a mammal. The cell may be *in vitro* (*e.g.* in cell culture) or *in vivo*. Preferred epithelial cells are from the nasopharynx. The cells are most preferably human cells.

Blocking the Neisseria–epithelium interaction

The invention provides a method for preventing the attachment of a Neisserial cell to an epithelial cell, wherein the ability of one or more App, ORF40 and/or NadA to bind to the epithelial cell is blocked.

The ability to bind may be blocked in various ways but, most conveniently, an antibody specific for App, ORF40 and/or NadA is used. The invention also provides antibody which is specific for App, ORF40 or NadA. This antibody preferably has an affinity for App, ORF40 and/or NadA of at least 10^{-7} M *e.g.* 10^{-8} M, 10^{-9} M, 10^{-10} M or tighter.

Antibodies for use in accordance with the invention may be polyclonal, but are preferably monoclonal. It will be appreciated that the term "antibody" includes whole antibodies (*e.g.* IgG, IgA *etc*), derivatives of whole antibodies which retain the antigen-binding sites (*e.g.* F_{ab}, $F_{ab'}$, $F_{(ab')2}$ *etc.*), single chain antibodies (*e.g.* sFv), chimeric antibodies, CDR-grafted antibodies, humanised antibodies, univalent antibodies, human monoclonal antibodies [*e.g.* Green (1999) *J Immunol Methods* 231:11-23; Kipriyanov & Little (1999) *Mol Biotechnol* 12:173-201 *etc.*] and the like. Humanised antibodies may be preferable to those which are fully human [*e.g.* Fletcher (2001) *Nature Biotechnology* 19:395-96].

As an alternative to using antibodies, antagonists of the interaction between App, ORF40 or NadA and its receptor on the epithelial cell may be used. As a further alternative, a soluble form of the epithelial cell receptor may be used as a decoy. These can be produced by removing the receptor's transmembrane and, optionally, cytoplasmic regions [*e.g.* EP-B2-0139417, EP-A-0609580 *etc.*].

The antibodies, antagonists and soluble receptors of the invention may be used as medicaments to prevent the attachment of a Neisserial cell to an epithelial cell.

Inhibiting expression of the Neisserial gene

The invention provides a method for preventing the attachment of a Neisserial cell to an epithelial cell, wherein protein expression from one or more of App, ORF40 and/or NadA is inhibited. The inhibition may be at the level of transcription and/or translation.

A preferred technique for inhibiting expression of the gene is antisense [*e.g.* Piddock (1998) *Curr Opin Microbiol* 1:502-8; Nielsen (2001) *Expert Opin Investig Drugs* 10:331-41; Good & Nielsen (1998) *Nature Biotechnol* 16:355-358; Rahman *et al.* (1991) *Antisense Res Dev* 1:319-327; *Methods in Enzymology* volumes 313 & 314; *Manual of Antisense Methodology* (eds. Hartmann & Endres); *Antisense Therapeutics* (ed. Agrawal) *etc.*]. Antibacterial antisense techniques are disclosed in, for example, international patent applications WO99/02673 and WO99/13893.

The invention also provides nucleic acid comprising a fragment of x or more nucleotides from nucleic acid which encodes App, ORF40 or NadA, wherein x is at least 8 (*e.g.* 8, 10, 12, 14, 16, 18, 20, 25, 30 or more). The nucleic acid will typically be single-stranded.

The nucleic acid is preferably of the formula 5'–$(N)_a$–(X)–$(N)_b$–3', wherein $0 \geq a \geq 15$, $0 \geq b \geq 15$, N is any nucleotide, and X is a fragment of a nucleic acid which encodes App, ORF40 or NadA. X preferably comprises at least 8 nucleotides (*e.g.* 8, 10, 12, 14, 16, 18, 20, 25, 30 or more). The values

of *a* and *b* may independently be 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 or 15. Each individual nucleotide N in the $-(N)_a-$ and $-(N)_b-$ portions of the nucleic acid may be the same or different. The length of the nucleic acid (*i.e.* a+b+length of X) is preferably less than 100 (*e.g.* less than 90, 80, 70, 60, 50, 40, 30 *etc.*).

It will be appreciated that the term "nucleic acid" includes DNA, RNA, DNA/RNA hybrids, DNA and RNA analogues such as those containing modified backbones (with modifications in the sugar and/or phosphates *e.g.* phosphorothioates, phosphoramidites *etc.*), and also peptide nucleic acids (PNA) and any other polymer comprising purine and pyrimidine bases or other natural, chemically or biochemically modified, non-natural, or derivatized nucleotide bases *etc.* Nucleic acid according to the invention can be prepared in many ways (*e.g.* by chemical synthesis, from genomic or cDNA libraries, from the organism itself *etc.*) and can take various forms (*e.g.* single stranded, double stranded, vectors, probes *etc.*).

The antisense nucleic acids of the invention may be used as medicaments to prevent the attachment of a Neisserial cell to an epithelial cell.

Knockout of the Neisserial gene

The invention provides a method for preventing the attachment of a Neisserial cell to an epithelial cell, wherein one or more of App, ORF40 and/or NadA is knocked out.

The invention also provides a *Neisseria* bacterium in which one or more of App, ORF40 and/or NadA has been knocked out.

Techniques for producing knockout bacteria are well known, and knockout *Neisseria* have been reported [*e.g.* Moe *et al.* (2001) *Infect. Immun.* 69:3762-3771; Seifert (1997) *Gene* 188:215-220; Zhu *et al.* (2000) *J.Bacteriol.* 182:439-447 *etc.*].

The knockout mutation may be situated in the coding region of the gene or may lie within its transcriptional control regions (*e.g.* within its promoter).

The knockout mutation will reduce the level of mRNA encoding App, ORF40 and/or NadA to <1% of that produced by the wild-type bacterium, preferably <0.5%, more preferably <0.1%, and most preferably to 0%.

The knockout mutants of the invention may be used as immunogenic compositions (*e.g.* as vaccines) to prevent Neisserial infection. Such a vaccine may include the mutant as a live attenuated bacterium.

Mutagenesis of the Neisserial gene

The invention provides a method for preventing the attachment of a Neisserial cell to an epithelial cell, wherein one or more of App, ORF40 and/or NadA has a mutation which inhibits its activity.

The invention also provides a mutant protein, wherein the mutant protein comprises the amino acid sequence of App, ORF40 and/or NadA, or a fragment thereof, but wherein one or more amino acids of said amino acid sequence is/are mutated (*e.g.* see below for App).

The amino acids which is/are mutated preferably result in the reduction or removal of an activity of App, ORF40 and/or NadA which is responsible directly or indirectly for adhesion to epithelial cells.

For example, the mutation may inhibit an enzymatic activity or may remove a binding site in the protein.

The invention also provides nucleic acid encoding this mutant protein.

The invention also provides a method for producing this nucleic acid, comprising the steps of: (a) providing source nucleic acid encoding App, ORF40 or NadA, and (b) performing mutagenesis (*e.g.* site-directed mutagenesis) on said source nucleic acid to provide nucleic acid encoding a mutant protein.

Mutation may involve deletion, substitution, and/or insertion, any of which may be involve one or more amino acids. As an alternative, the mutation may involve truncation.

Mutagenesis of virulence factors is a well-established science for many bacteria [*e.g.* toxin mutagenesis described in WO93/13202; Rappuoli & Pizza, Chapter 1 of *Sourcebook of Bacterial Protein Toxins* (ISBN 0-12-053078-3); Pizza *et al.* (2001) *Vaccine* 19:2534-41; Alape-Giron *et al.* (2000) *Eur J Biochem* 267:5191-5197; Kitten *et al.* (2000) *Infect Immun* 68:4441-4451; Gubba *et al.* (2000) *Infect Immun* 68:3716-3719; Boulnois *et al.* (1991) *Mol Microbiol* 5:2611-2616 *etc.*] including *Neisseria* [*e.g.* Power *et al.* (2000) *Microbiology* 146:967-979; Forest *et al.* (1999) *Mol Microbiol* 31:743-752; Cornelissen *et al.* (1998) *Mol Microbiol* 27:611-616; Lee *et al.* (1995) *Infect Immun* 63:2508-2515; Robertson *et al.* (1993) *Mol Microbiol* 8:891-901 *etc.*].

Mutagenesis may be specifically targeted to nucleic acid encoding App, ORF40 and/or NadA. Alternatively, mutagenesis may be global or random (*e.g.* by irradiation, chemical mutagenesis *etc.*), which will typically be followed by screening bacteria for those in which a mutation has been introduced into App, ORF40 and/or NadA. Such screening may be by hybridisation assays (*e.g.* Southern or Northern blots *etc.*), primer-based amplification (*e.g.* PCR), sequencing, proteomics, aberrant SDS-PAGE gel migration *etc.*

The mutant proteins and nucleic acids of the invention may be used as immunogenic compositions (*e.g.* as vaccines) to prevent Neisserial infection.

Screening methods

The invention also provides methods for screening compounds to identify those (antagonists) which inhibit the binding of a Neisserial cell to an epithelial cell.

Potential antagonists for screening include small organic molecules, peptides, peptoids, polypeptides, lipids, metals, nucleotides, nucleosides, polyamines, antibodies, and derivatives thereof. Small organic molecules have a molecular weight between 50 and about 2,500 daltons, and most preferably in the range 200-800 daltons. Complex mixtures of substances, such as extracts containing natural products, compound libraries or the products of mixed combinatorial syntheses also contain potential antagonists.

Typically, App, ORF40 and/or NadA protein is incubated with an epithelial cell and a test compound, and the mixture is then tested to see if the interaction between the protein and the epithelial cell has been inhibited.

Inhibition will, of course, be determined relative to a standard (*e.g.* the native protein/cell interaction). Preferably, the standard is a control value measured in the absence of the test compound. It will be appreciated that the standard may have been determined before performing the method, or may be determined during or after the method has been performed. It may also be an absolute standard.

The protein, cell and compound may be mixed in any order.

For preferred high-throughput screening methods, all the biochemical steps for this assay are performed in a single solution in, for instance, a test tube or microtitre plate, and the test compounds are analysed initially at a single compound concentration. For the purposes of high throughput screening, the experimental conditions are adjusted to achieve a proportion of test compounds identified as "positive" compounds from amongst the total compounds screened.

Other methods which may be used include, for example, reverse two hybrid screening [*e.g.* Vidal & Endoh (1999) *TIBTECH* 17:374-381] in which the inhibition of the Neisseria:receptor interaction is reported as a failure to activate transcription.

The method may also simply involve incubating one or more test compound(s) with App, ORF40 and/or NadA and determining if they interact. Compounds that interact with the protein can then be tested for their ability to block an interaction between the protein and an epithelial cell.

The invention also provides a compound identified using these methods. These can be used to treat or prevent Neisserial infection. The compound preferably has an affinity for App, ORF40 and/or NadA of at least 10^{-7} M *e.g.* 10^{-8} M, 10^{-9} M, 10^{-10} M or tighter.

The invention also provides a composition comprising (a) an *E.coli* bacterium which expresses App and/or ORF40 (and, optionally, NadA) and (b) an epithelial cell (*e.g.* a human epithelial cell).

Expression in outer membrane vesicles (OMVs)

International patent application WO01/52885 discloses that the addition of further defined components to OMV vaccines significantly broadens their efficacy.

The preparation of OMVs from NmB is well-known in the art. Methods for obtaining suitable preparations are disclosed in, for instance: Claassen *et al.* [*Vaccine* (1996) 14:1001-1008]; Cartwright *et al.* [*Vaccine* (1999) 17:2612-2619]; Peeters *et al.* [*Vaccine* (1996) 14:1009-1015]; Fu *et al.* [*Biotechnology NY* (1995) 12:170-74]; Davies *et al.* [*J.Immunol.Meth.* (1990) 134:215-225]; Saunders *et al.* [*Infect. Immun.* (1999) 67:113-119]; Draabick *et al.* [*Vaccine* (2000) 18:160-172]; Moreno *et al.* [*Infect. Immun.* (1985) 47:527-533]; Milagres *et al.* [*Infect. Immun.* (1994) 62:4419-4424]; Naess *et al.* [*Infect. Immun.* (1998) 66:959-965]; Rosenqvist *et al.* [*Dev.Biol.Stand.* (1998) 92:323-333]; Haneberg *et al.* [*Infect. Immunn.* (1998) 66:1334-41]; Andersen *et al.* [*Vaccine* (1997) 15:1225-34]; Bjune *et al.* [*Lancet* (1991) 338:1093-96] *etc.*

It has now been found that OMVs prepared from *E.coli* which express a heterologous *Neisseria* gene can give better results in standard immunogenicity tests than the antigens in purified form.

The invention therefore provides a method for preparing an OMV from a non-Neisserial host cell, characterised in that said cell expresses a gene encoding App, ORF40 or NadA protein.

The invention also provides (a) OMVs obtainable by this process, and (b) an outer membrane vesicle from a non-Neisserial host cell, characterised in that said cell expresses a gene encoding App, ORF40 or NadA protein.

The non-Neisserial host cell is preferably a bacterium and is most preferably *E.coli*.

More generally, the invention provides a method for preparing an OMV from a non-Neisserial host cell, characterised in that said cell expresses a gene encoding one or more of the following proteins:

(A) Even SEQ IDs 2-892 from WO99/24578;

(B) Even SEQ IDs 2-90 from WO99/36544;

(C) Even SEQ IDs 2-3020 from WO99/57280;

(D) Even SEQ IDs 3040-3114 from WO99/57280;

(E) SEQ IDs 3115-3241 from WO99/57280;

(F) The 2160 proteins NMB0001 to NMB2160 from Tettelin *et al.* [*supra*];

(G) A protein comprising the amino acid sequence of one or more of (A) to (F);

(H) A protein sharing sequence identity with the amino acid sequence of one or more of (A) to (F); and

(I) A protein comprising a fragment of one or more of (A) to (F).

Similarly, the invention also provides (a) OMVs obtainable by this process, and (b) an outer membrane vesicle from a non-Neisserial host cell, characterised in that said cell expresses a gene encoding one or more of proteins (A) to (I) described above.

The degree of 'sequence identity' referred to in (H) is preferably greater than 50% (*eg.* 60%, 70%, 80%, 90%, 95%, 99% or more) and this includes mutants and allelic variants

The 'fragment' referred to in (I) should comprise at least *n* consecutive amino acids from one or more of (A) to (F) and, depending on the particular sequence, *n* is 7 or more (*eg.* 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100 or more). Preferably the fragment comprises an epitope from one or more of (A) to (F). Preferred fragments are those disclosed in WO00/71574 and WO01/04316.

Preferred proteins for (A) to (F) are found in *N.meningitidis* serogroup B.

Mutants of App

Amino acid 267 of SEQ ID 650 of WO99/24578 (SEQ ID 32 herein) is a serine. App is believed to be a serine protease and this serine is believed to be a catalytic residue at its active site. It will be appreciated that standard sequence alignment techniques will reveal the amino acid corresponding to this Ser-267 for any other App sequence (*e.g.* Ser-260 in SEQ ID 652 of WO99/24578, Ser-267 in SEQ ID 654 *etc.*).

The invention provides a protein comprising the amino acid sequence of App, except that one or more of amino acids Ser-267, Asp-158 and His-115 (numbered according to SEQ ID 32) is/are

mutated. The mutation may be a deletion, an insertion or, preferably, a substitution. The substitution is preferably with one of the 19 other naturally-occurring amino acids and is more preferably with glycine, alanine, tyrosine or lysine.

App is believed to cleaved at a site between amino acids 1063 and 1171 (numbered according to SEQ ID 32). It will be appreciated that standard sequence alignment techniques will reveal the amino acids corresponding to these two residues for any other App sequence.

The invention provides a protein comprising the amino acid sequence of App, except that one or more amino acid(s) between Ser-1064 and Arg-1171 (numbered according to SEQ ID 32) is mutated. The mutation may be a deletion, an insertion, truncation or, preferably, a substitution. The substitution is preferably with one of the 19 other naturally-occurring amino acids. The residue which is mutated is preferably S-1064, D-1065, K-1066, L-1067, G-1068, K-1069, A-1070, E-1071, A-1072, K-1073, K-1074, Q-1075, A-1076, E-1077, K-1078, D-1079, N-1080, A-1081, Q-1082, S-1083, L-1084, D-1085, A-1086, L-1087, I-1088, A-1089, A-1090, G-1091, R-1092, D-1093, A-1094, V-1095, E-1096, K-1097, T-1098, E-1099, S-1100, V-1101, A-1102, E-1103, P-1104, A-1105, R-1106, Q-1107, A-1108, G-1109, G-1110, E-1111, N-1112, V-1113, G-1114, I-1115, M-1116, Q-1117, A-1118, E-1119, E-1120, E-1121, K-1122, K-1123, R-1124, V-1125, Q-1126, A-1127, D-1128, K-1129, D-1130, T-1131, A-1132, L-1133, A-1134, K-1135, Q-1136, R-1137, E-1138, A-1139, E-1140, T-1141, R-1142, P-1143, A-1144, T-1145, T-1146, A-1147, F-1148, P-1149, R-1150, A-1151, R-1152, R-1153, A-1154, R-1155, R-1156, D-1157, L-1158, P-1159, Q-1160, L-1161, Q-1162, P-1163, Q-1164, P-1165, Q-1166, P-1167, Q-1168, P-1169, Q-1170 and/or R-1171.

App is alternatively believed to cleaved at amino acid 956 and/or amino acid 1178 (numbered according to SEQ ID 32). It will be appreciated that standard sequence alignment techniques will reveal the amino acids corresponding to these residues for any other App sequence.

The invention provides a protein comprising the amino acid sequence of App, except that one or more of amino acids Phe-956, Asn-957, Ala-1178 & Asn-1179 (numbered according to SEQ ID 32) is mutated. The mutation may be a deletion, an insertion, truncation or, preferably, a substitution. The substitution is preferably with one of the 19 other naturally-occurring amino acids.

The invention also provides nucleic acid encoding these mutant proteins.

The invention also provides a method for producing this nucleic acid, comprising the steps of: (a) providing source nucleic acid encoding App, ORF40 or NadA, and (b) performing mutagenesis (*e.g.* site-directed mutagenesis) on said source nucleic acid to provide nucleic acid encoding a mutant protein.

The invention provides mature App.

The invention also provides a protein comprising the amino acid sequence of a processed App, wherein said processed App does not comprise the C-terminus domain which is downstream of an autoproteloytic cleavage site in full-length App. For example, based on SEQ ID 32 as full-length App, the invention provides SEQ IDs 33 to 36. C-terminus domains which may be removed during autoproteolysis are SEQ IDs 38 and 39.

The invention also provides a protein comprising the amino acid sequence of a processed App, wherein the C-terminus of said processed App is Phe-956 (numbered according to SEQ ID 32). For example, the invention provides SEQ IDs 33 and 35. The amino acid corresponding to Phe-956 in other App sequences can be identified by standard sequence alignment techniques.

The invention also provides a protein comprising the amino acid sequence of a processed App, wherein the C-terminus of said processed App is Ala-1178 (numbered according to SEQ ID 32). For example, the invention provides SEQ IDs 34 and 36. The amino acid corresponding to Ala-1178 in other App sequences can be identified by standard sequence alignment techniques.

The invention also provides a protein comprising the amino acid sequence of a processed App, wherein said processed App does not comprise SEQ ID 37, 38 or 39.

The invention also provides a protein comprising an amino acid sequence selected from the group consisting of SEQ IDs 33, 34, 35, 36, 37, 38 & 39.

The invention also provides a protein comprising an amino acid sequence with at least *p*% sequence identity to one or more of SEQ IDs 33, 34, 35, 36, 37, 38 & 39. Depending on the particular sequence, the value of *p* is preferably 50 or more (*e.g.* 60, 70, 80, 90, 95, 99 or more). These proteins include homologs, orthologs, allelic variants and functional mutants. Typically, 50% identity or more between two proteins is considered to be an indication of functional equivalence. Identity between proteins is preferably determined by the Smith-Waterman homology search algorithm as implemented in the MPSRCH program (Oxford Molecular), using an affine gap search with parameters *gap open penalty=12* and *gap extension penalty=1*.

The invention further provides proteins comprising a fragment of one or more of SEQ IDs 33, 34, 35, 36, 37, 38 & 39. The fragments should comprise at least *q* consecutive amino acids from the sequences and, depending on the particular sequence, *q* is 7 or more (*e.g.* 8, 10, 12, 14, 16, 18, 20, 30, 40, 50, 60, 70, 80, 90, 100 or more). Preferably the fragments comprise one or more epitopes from the sequence.

The invention also provides nucleic acid encoding these proteins of the invention.

Alleles of NadA

The invention provides a protein comprising the amino acid sequence of one or more of SEQ IDs 1 to 14.

The invention also provides a protein comprising an amino acid sequence having at least *x*% sequence identity to one or more of SEQ IDs 1 to 14. The value of *x* is at least 50% (*e.g.* 60%, 70%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, 99.5% or more). This includes variants *e.g.* allelic variants, homologs, orthologs, paralogs, mutants, *etc.*

A preferred allele of NadA for use with the present invention is SEQ ID 3 (or SEQ ID 6).

The invention also provides a protein comprising a fragment of one or more of SEQ IDs 1 to 14. These should comprise at least *n* consecutive nucleotides from one or more of SEQ IDs 1 to 14, wherein *n* is 6 or more (*e.g.* 7, 8, 9, 10, 11, 12, 14, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100,

150, 200, 250, 300, 350 or more). The fragment may comprise a sequence which is common to SEQ IDs 1 to 14, or may comprise a sequence which is not common to SEQ IDs 1 to 14.

Preferred fragments comprise one or more epitopes from SEQ IDs 1 to 14. Other preferred fragments are (a) the N-terminal leader peptides of SEQ IDs 1 to 14, (b) SEQ IDs 1 to 14, but without *k* N-terminal amino acid residue(s), wherein *k* is 1 or more (*e.g.* 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40, 50 *etc.*), and (c) SEQ IDs 1 to 14, but without *l* C-terminal amino acid residue(s), wherein *l* is 1 or more (*e.g.* 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40, 50 *etc.*). Preferred fragments fall within both (b) and (c) *i.e.* truncation at both C- and N- termini.

Preferred fragments within category (b) lack the N-terminal leader peptide. For SEQ IDs 1, 2, 3, 7, 9, 11 & 13 the value of *k* is thus 23; for SEQ IDs 4, 5, 6, 8, 10, 12 & 14 the value of *k* is 25. The leader peptide may be replaced with the leader peptide from another protein, by another protein (*i.e.* to form a fusion protein) or by an alternative N-terminus sequence to allow efficient expression.

Preferred fragments within category (c) lack the C-terminal membrane anchor. The value of *l* is thus 54. Minor variants of this C-terminal deletion may be used (*e.g.* where *l* is 45, 46, 47, 48, 49, 50, 51, 52, 53, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66).

Proteins with the N-terminus sequence MKH or MQH are preferred to those with N-terminus sequence MSM.

The protein of the invention may include the heptad sequence $(AA_1AA_2AA_3AA_4AA_5AA_6AA_7)_r$ wherein: AA_1 is Leu, Ile, Val or Met; each of AA_2 AA_3 AA_4 AA_5 AA_6 and AA_7 may independently be any amino acid; *r* is an integer of 1 or more (*e.g.* 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 *etc.*). Where *r* is 2 or more, the meaning of each AA_1 AA_2 AA_3 AA_4 AA_5 AA_6 and AA_7 may be the same or different in each of the *r* heptad repeats. The heptad(s) can form a leucine-zipper domain.

Proteins of the invention can be prepared in many ways *e.g.* by chemical synthesis (at least in part), by digesting longer polypeptides using proteases, by translation from RNA, by purification from cell culture (*e.g.* from recombinant expression), from the organism itself (*e.g.* isolation from prostate tissue), from a cell line source, *etc.*

Proteins of the invention can be prepared in various forms *e.g.* native, fusions, glycosylated, non-glycosylated, lipidated, non-lipidated *etc.*

The protein is preferably in the form of an oligomer.

Proteins of the invention may be attached or immobilised to a solid support.

Proteins of the invention may comprise a detectable label *e.g.* a radioactive label, a fluorescent label, or a biotin label. This is particularly useful in immunoassay techniques.

Proteins of the invention are preferably in isolated or substantially isolated form.

In general, the proteins of the invention are provided in a non-naturally occurring environment *e.g.* they are separated from their naturally-occurring environment. In certain embodiments, the subject protein is present in a composition that is enriched for the protein as compared to a control. As such,

purified protein is provided, whereby purified is meant that the protein is present in a composition that is substantially free of other expressed proteins, where by substantially free is meant that less than 90%, usually less than 60% and more usually less than 50% of the composition is made up of other expressed proteins.

The term "protein" refers to amino acid polymers of any length. The polymer may be linear or branched, it may comprise modified amino acids, and it may be interrupted by non-amino acids. The terms also encompass an amino acid polymer that has been modified naturally or by intervention; for example, disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, or any other manipulation or modification, such as conjugation with a labeling component. Also included within the definition are, for exampleproteins containing one or more analogs of an amino acid (including, for example, unnatural amino acids, *etc.*), as well as other modifications known in the art. Proteins can occur as single chains or associated chains.

Mutants can include amino acid substitutions, additions or deletions. The amino acid substitutions can be conservative amino acid substitutions or substitutions to eliminate non-essential amino acids, such as to alter a glycosylation site, a phosphorylation site or an acetylation site, or to minimize misfolding by substitution or deletion of one or more cysteine residues that are not necessary for function. Conservative amino acid substitutions are those that preserve the general charge, hydrophobicity/hydrophilicity, and/or steric bulk of the amino acid substituted. Variants can be designed so as to retain or have enhanced biological activity of a particular region of the polypeptide (*e.g.* a functional domain and/or, where the polypeptide is a member of a polypeptide family, a region associated with a consensus sequence). Selection of amino acid alterations for production of variants can be based upon the accessibility (interior *vs.* exterior) of the amino acid, the thermostability of the variant polypeptide, desired disulfide bridges, desired metal binding sites *etc.*

The invention also provides nucleic acid encoding a protein of the invention as defined above. The invention also provides nucleic acid comprising a fragment of at least *n* consecutive nucleotides from said nucleic acid, wherein *n* is 10 or more (*e.g.* 12, 14, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 500 or more).

Furthermore, the invention provides nucleic acid which can hybridise to nucleic acid encoding a protein of the invention, preferably under "high stringency" conditions (*eg.* 65°C in a 0.1xSSC, 0.5% SDS solution).

Nucleic acids of the invention can be used in hybridisation reactions (*e.g.* Northern or Southern blots, or in nucleic acid microarrays or 'gene chips') and amplification reactions (*e.g.* PCR, SDA, SSSR, LCR, TMA, NASBA, *etc.*) and other nucleic acid techniques.

Nucleic acids of the invention can be prepared in many ways *e.g.* by chemical synthesis in whole or part, by digesting longer polynucleotides using nucleases (*e.g.* restriction enzymes), from genomic or cDNA libraries, from the bacterium itself, *etc.*

Nucleic acids of the invention can take various forms *e.g.* single-stranded, double-stranded, vectors, primers, probes, labelled, unlabelled, *etc.*

Nucleic acids of the invention are preferably in isolated or substantially isolated form.

The invention includes nucleic acid comprising sequences complementary to those described above *e.g.* for antisense or probing, or for use as primers.

The term "nucleic acid" includes DNA and RNA, and also their analogues, such as those containing modified backbones, and also peptide nucleic acids (PNA) *etc.*

Nucleic acid according to the invention may be labelled *e.g.* with a radioactive or fluorescent label. This is particularly useful where the nucleic acid is to be used in nucleic acid detection techniques *e.g.* where the nucleic acid is a primer or as a probe for use in techniques such as PCR, LCR, TMA, NASBA, *etc.*

The invention also provides vectors comprising nucleotide sequences of the invention (*e.g.* cloning or expression vectors, such as those suitable for nucleic acid immunisation) and host cells transformed with such vectors.

Immunisation

The invention provides an immunogenic composition comprising (a) a Neisserial NadA protein and/or (b) nucleic acid encoding a NadA protein.

The invention also provides a method for raising an antibody response in a mammal, comprising administering an immunogenic composition of the invention to the mammal. The antibody response is preferably a protective antibody response. The protective antibody preferably blocks the attachment of NadA and/or App to epithelial cells.

The invention also provides a method for protecting a mammal against a Neisserial infection, comprising administering to the mammal an immunogenic composition of the invention.

The invention also provides Neisserial NadA protein for use as a medicament.

The invention also provides the use of a NadA protein in the manufacture of a medicament for preventing Neisserial infection in a mammal

The invention also provides the use of nucleic acid encoding a NadA protein in the manufacture of a medicament for preventing Neisserial infection in a mammal.

The mammal is preferably a human. The human may be an adult or, preferably, a child.

The NadA protein is preferably a *N.meningitidis* NadA. It preferably comprises the amino acid sequence of one or more of SEQ IDs 1 to 14, or an amino acid sequence having sequence identity thereto or comprising a fragment thereof (see above). The NadA protein is preferably in the form of an oligomer (*e.g.* a dimer, trimer, tetramer or higher). Within SEQ IDs 1 to 14, SEQ IDs 1 to 12 are preferred, as antibodies against these NadA proteins are bactericidal across the various hypervirulent alleles. Where an immune response against a non-hypervirulent $NadA^+$ strain is desired, however, SEQ IDs 13 & 14 are preferred. Of course, NadA mixtures are also possible, particularly mixtures containing more than one NadA allele.

Immunogenic compositions of the invention may be used therapeutically (*i.e.* to treat an existing infection) or prophylactically (*i.e.* to prevent future infection).

The uses and methods of the invention are particularly useful for treating/protecting against infections of *Neisseria meningitidis*, including serogroups A, B, and C. They are particularly useful against strains of *N.meningitidis* from hypervirulent lineages ET-5, EY-37 and cluster A4.

The uses and methods are particularly useful for preventing/treating diseases including, but not limited to, meningitis (particularly bacterial meningitis) and bacteremia.

Efficacy of therapeutic treatment can be tested by monitoring Neisserial infection after administration of the composition of the invention. Efficacy of prophylactic treatment can be tested by monitoring immune responses against NadA after administration of the composition.

The composition of the invention may additionally comprise an antigen which, when administered to a mammal, elicits an immune response which is protective against a lineage III strain of *N.meningitidis*.

Compositions of the invention will generally be administered directly to a patient. Direct delivery may be accomplished by parenteral injection (*e.g.* subcutaneously, intraperitoneally, intravenously, intramuscularly, or to the interstitial space of a tissue), or by rectal, oral, vaginal, topical, transdermal, intranasal, ocular, aural, or pulmonary administration.

The invention may be used to elicit systemic and/or mucosal immunity.

Dosage treatment can be a single dose schedule or a multiple dose schedule.

The immunogenic composition of the invention will generally include a pharmaceutically acceptable carrier, which can be any substance that does not itself induce the production of antibodies harmful to the patient receiving the composition, and which can be administered without undue toxicity. Suitable carriers can be large, slowly-metabolised macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, and inactive virus particles. Such carriers are well known to those of ordinary skill in the art. Pharmaceutically acceptable carriers can include liquids such as water, saline, glycerol and ethanol. Auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, can also be present in such vehicles. Liposomes are suitable carriers. A thorough discussion of pharmaceutical carriers is available in Gennaro (2000) *Remington: The Science and Practice of Pharmacy.* 20th edition, ISBN: 0683306472.

Neisserial infections affect various areas of the body and so the compositions of the invention may be prepared in various forms. For example, the compositions may be prepared as injectables, either as liquid solutions or suspensions. Solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection can also be prepared. The composition may be prepared for topical administration *e.g.* as an ointment, cream or powder. The composition be prepared for oral administration *e.g.* as a tablet or capsule, or as a syrup (optionally flavoured). The composition may be prepared for pulmonary administration *e.g.* as an inhaler, using a fine powder or a spray. The composition may be

prepared as a suppository or pessary. The composition may be prepared for nasal, aural or ocular administration *e.g.* as drops.

The composition is preferably sterile. It is preferably pyrogen-free. It is preferably buffered *e.g.* at between pH 6 and pH 8, generally around pH 7.

Immunogenic compositions comprise an immunologically effective amount of immunogen, as well as any other of other specified components, as needed. By 'immunologically effective amount', it is meant that the administration of that amount to an individual, either in a single dose or as part of a series, is effective for treatment or prevention. This amount varies depending upon the health and physical condition of the individual to be treated, age, the taxonomic group of individual to be treated (*e.g.* non-human primate, primate, *etc.*), the capacity of the individual's immune system to synthesise antibodies, the degree of protection desired, the formulation of the vaccine, the treating doctor's assessment of the medical situation, and other relevant factors. It is expected that the amount will fall in a relatively broad range that can be determined through routine trials. Dosage treatment may be a single dose schedule or a multiple dose schedule (*e.g.* including booster doses). The composition may be administered in conjunction with other immunoregulatory agents.

The immunogenic composition may include an adjuvant. Preferred adjuvants to enhance effectiveness of the composition include, but are not limited to: (A) aluminium compounds (*e.g.* an aluminium hydroxide such as oxyhydroxide, or an aluminium phosphatesuch as hydroxyphosphate or orthophosphate, aluminium sulphate *etc.*), or mixtures of different aluminium compounds, with the compounds taking any suitable form (*e.g.* gel, crystalline, amorphous *etc.*), and with adsorption being preferred; (B) MF59 (5% Squalene, 0.5% Tween 80, and 0.5% Span 85, formulated into submicron particles using a microfluidizer); (C) liposomes; (D) ISCOMs, which may be devoid of additional detergent; (E) SAF, containing 10% Squalane, 0.4% Tween 80, 5% pluronic-block polymer L121, and thr-MDP, either microfluidized into a submicron emulsion or vortexed to generate a larger particle size emulsion; (F) Ribi™ adjuvant system (RAS), (Ribi Immunochem) containing 2% Squalene, 0.2% Tween 80, and one or more bacterial cell wall components from the group consisting of monophosphorylipid A (MPL), trehalose dimycolate (TDM), and cell wall skeleton (CWS), preferably MPL + CWS (Detox™); (G) saponin adjuvants, such as QuilA or QS21, also known as Stimulon™; (H) chitosan; (I) complete Freund's adjuvant (CFA) and incomplete Freund's adjuvant (IFA); (J) cytokines, such as interleukins (*e.g.* IL-1, IL-2, IL-4, IL-5, IL-6, IL-7, IL-12, *etc.*), interferons (*e.g.* interferon-γ), macrophage colony stimulating factor, tumor necrosis factor, *etc.*; (K) microparticles (*i.e.* a particle of ~100nm to ~150μm in diameter, more preferably ~200nm to ~30μm in diameter, and most preferably ~500nm to ~10μm in diameter) formed from materials that are biodegradable and non-toxic (*e.g.* a poly(α-hydroxy acid), a polyhydroxybutyric acid, a polyorthoester, a polyanhydride, a polycaprolactone *etc.*); (L) monophosphoryl lipid A (MPL) or 3-O-deacylated MPL (3dMPL); (M) combinations of 3dMPL with, for example, QS21 and/or oil-in-water emulsions; (N) oligonucleotides comprising CpG motifs *i.e.* containing at least one CG dinucleotide, with 5-methylcytosine optionally being used in place of cytosine; (O) a polyoxyethylene ether or a polyoxyethylene ester; (P) a polyoxyethylene sorbitan ester surfactant in

combination with an octoxynol or a polyoxyethylene alkyl ether or ester surfactant in combination with at least one additional non-ionic surfactant such as an octoxynol; (Q) an immunostimulatory oligonucleotide (*e.g.* a CpG oligonucleotide) and a saponin; (R) an immunostimulant and a particle of metal salt; (S) a saponin and an oil-in-water emulsion; (T) a saponin (*e.g.* QS21) + 3dMPL + IL-12 (optionally + a sterol); (U) *E.coli* heat-labile enterotoxin ("LT"), or detoxified mutants thereof, such as the K63 or R72 mutants; (V) cholera toxin ("CT"), or detoxified mutants thereof; (W) microparticles (*i.e.* a particle of ~100nm to ~150μm in diameter, more preferably ~200nm to ~30μm in diameter, and most preferably ~500nm to ~10μm in diameter) formed from materials that are biodegradable and non-toxic (*e.g.* a poly(α-hydroxy acid) such as poly(lactide-co-glycolide), a polyhydroxybutyric acid, a polyorthoester, a polyanhydride, a polycaprolactone *etc.*); and (X) other substances that act as immunostimulating agents to enhance the effectiveness of the composition. Aluminium salts (aluminium phosphates and particularly hydroxyphosphates, and/or hydroxides and particularly oxyhydroxide) and MF59 are preferred adjuvants for parenteral immunisation. Toxin mutants are preferred mucosal adjuvants.

Muramyl peptides include N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acetyl-normuramyl-L-alanyl-D-isoglutamine (nor-MDP), N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-*sn*-glycero-3-hydroxyphosphoryloxy)-ethylamine MTP-PE), *etc.*

Compositions of the invention may comprise antigens (*e.g.* protective antigens against *N.meningitidis* or against other organisms) in addition to NadA *e.g.* DTP antigens, Hib antigen *etc.*

Immunogenic compositions of the invention may be used therapeutically (*i.e.* to treat an existing infection) or prophylactically (*i.e.* to prevent future infection). Therapeutic immunisation is particularly useful for treating *Candida* infection in immunocompromised subjects.

As an alternative to using proteins antigens in the immunogenic compositions of the invention, nucleic acid (preferably DNA *e.g.* in the form of a plasmid) encoding the antigen may be used.

Disclaimers

The invention preferably excludes: (a) amino acid and nucleic acid sequences available in public sequence databases (*e.g.* GenBank or GENESEQ) prior to 26th July 2002 and, more preferably, prior to 27th July 2001; (b) amino acid and nucleic acid sequences disclosed in patent applications having a filing date or, where applicable, a priority date prior to 26th July 2002 and, more preferably, prior to 27th July 2001. In particular, SEQ ID entries in the following patent applications may be excluded: WO99/24578; WO99/36544; WO99/57280; WO00/22430; WO00/66741; WO00/66791; WO00/71574; WO00/71725; WO01/04316; WO01/31019; WO01/37863; WO01/38350; WO01/52885; WO01/64920; WO01/64922.

Definitions

The term "comprising" means "including" as well as "consisting" *e.g.* a composition "comprising" X may consist exclusively of X or may include something additional *e.g.* X + Y.

BRIEF DESCRIPTION OF DRAWINGS

Figures 1 to 3 show expression data for (1) ORF40 (2) App (3) NadA.

Figures 4 to 6 show FACS analysis of proteins involved in adhesion to human cells. In Figures 4 and 5 (Figure 6), the data are for, from left to right, ORF40 (▲), App (●), NadA (♦) and GNA2132 (■).

Figures 7 and 8 show homologies of (7) ORF40 and (8) App.

Figure 9 shows an alignment of NadA alleles, and figure 10 shows the relationship of alleles 1 to 3.

Figure 11 shows predicted secondary structure for NadA.

Figure 12 shows analysis of sequences upstream and downstream of NadA.

Figure 13 shows PCR analysis of NadA expression in different strains of *N.meningitidis*.

Figure 14 shows immunoblot analysis of NadA expression in different strains of *N.meningitidis*.

Figure 15 shows variation of NadA expression with culture time.

Figure 16 shows NadA FACS of isogenic capsulated and non-capsulated *N.meningitidis* cells.

Figure 17 shows immunofluorescence results obtained using anti-NadA against Chang cells (17A to 17C) or HeLa cells (17D).

Figure 18 shows immunofluorescence results obtained using anti-NadA against Chang cells after incubation at (A) 37°C or (B) 4°C.

Figure 19 shows immunofluorescence results for Chang cells treated with saponin.

Figure 20 shows immunofluorescence results obtained using monocytes.

Figure 21 shows immunofluorescence results obtained using macrophages.

Figure 22 shows IL-α secretion by monocytes in response to NadA treatment.

Figure 23 shows the effect of anti-CD14 on IL-α secretion by monocytes.

Figure 24 shows immunofluorescence results obtained using anti-NadA against *E.coli* transformed to express NadA.

Figure 25 shows staining of the transformed *E.coli* using (A) anti-NadA (B) anti-*E.coli* or (C) both.

Figure 26 is a schematic representation of App features. The N-terminal leader peptide, the passenger domain and the C-terminal β-domain are indicated. The positions of the serine protease active site, the ATP/GTP binding site, the two Arginine-rich sites and the Proline-rich region are shown. In BOX 1, cleavage sites are shown. In BOX 2 a comparison of known proteolytic sites of different autotransporters is shown and a consensus signature is derived. Arrows identify the cleavages; X = any amino acid; hyd = hydrophobic residues; (A,S) = Alanine or Serine.

Figure 27 is a schematic representation of the constructs used for studying App.

Figure 28 shows a western blot of outer membrane and extracellular proteins in *E.coli*.

Figure 29 shows FACS analysis of outer membrane and extracellular proteins in *E.coli*.

Figure 30 shows immunofluorescence of outer membrane and extracellular proteins in *E.coli*.

Figure 31 shows total *E.coli* proteins analysed by SDS-PAGE.

Figure 32 shows an immunoblot of crude precipitated culture supernatants using mouse antiserum against App-his.

Figure 33 shows FACS adhesion data using rabbit antiserum against *E.coli*. Percentages of cells positive to adhesion are shown near the fluorescence profiles.

Figure 34 shows immunofluorescence microscopy data showing bacterial adherence and aggregation.

Figure 35 shows concentration-dependent binding of App-His (♦), Appα-His (■) and NMB2132 (▲) expressed as net Mean Fluorescence Intensity (MFI).

Figure 36 shows the effect on binding of App-His (100μg/ml) of pre-incubation with pronase (left-hand columns) or phospholipase A2 (right-hand columns) with increasing concentration of enzyme. Pronase was tested at 0, 250, 500, 1000 μg/ml; phosholipase A2 was tested at 0, 50, 200, 800 μg/ml.

Figure 37 is a comparison of cellular binding specificity of App-His protein at 100, 25 or 6.25 μg/ml against various different cells.

Figure 38 shows association of wild-type or App-knockout *N.meningitidis* MC58 bacteria.

Figure 39 shows a western blot analysis of total lysates from *N.meningitidis* MC58 harvested at 0.5 or 0.8 OD_{620nm}. Lanes 1 & 3 show wild-type MC58 and lanes 2 & 4 show the App knockout.

Figure 40 shows a western blot analysis of supernatants in parallel to figure 39.

MODES FOR CARRYING OUT THE INVENTION

NadA homology

NadA shows homology to (a) YadA of enteropathogenic *Yersinia,* a non-pilus associated adhesin implicated in virulence [Cornelis (1998) *Microbiol. Mol. Biol. Rev.* 62:1315-1352.] and (b) UspA2 of *Moraxella catarrhalis,* a protein involved in serum resistance and a protective antigen [Chen *et al.* (1999) *Infect. Immun.* 67:1310-1316.]. Sequence similarity is mainly clustered in the carboxyl terminal region (56-63% identity in the last 70 amino acids). Outside this region the level of identity drops to 23-25%.

YadA and UspA2 have been identified as adhesins [Hoiczyk *et al.* (2000) *EMBO J* 19:5989-5999]. Both proteins form very stable and difficult-to-dissociate high molecular weight oligomers (150-200 kDa) anchored to the outer membrane. NadA has also been found to form very stable high molecular weight aggregates on the outer membrane of meningococcus.

The amino acid sequence of NadA was analysed [Nielsen *et al.* (1997) *Protein Engineering* 10:1-6; Levin & Garner (1988) *Biochim. Biophys. Acta* 955:283-295; Berger *et al.* (1995) *PNAS USA* 92:8259-8263; Bornberg-Bauer *et al.* (1998) *Nucleic Acids Res.* 26:2740-2746]. Secondary structure analysis is shown in Figure 11. The globular N-terminus and amphipathic C-terminus are indicated, as are the positions of the leader peptide (LP) and a membrane anchor. The carboxyl-terminal region (aa 310-362) has a predicted amphipatic β-structure (β-strands shown in black) and a terminal aromatic amino acid, which are typical features of outer membrane anchoring domains. The amino terminal region (aa 23-90) has no defined secondary structure, but the rest of the protein has mainly α-helix propensity (84.6%). Within this region, residues 90-146 and 183-288 have high probability of forming coiled coils. In addition, residues 122-143 contain four leucine residues in the "a" positions of the heptad repeats (L-x(6)-L-x(6)-L-x(6)-L) that may form a leucine zipper domain (●●●). It is

known that both coiled coils and leucine zipper sequences are involved in dimerization and may mediate oligomerisation of monomers via association of two or more alpha helices.

Even though primary structure similarity between NadA, YadA and UspA2 is clustered at the C-terminus, therefore, the overall similarity between the three proteins is conserved at secondary structure level. Putative leucine zippers are present in both NadA and UspA2. NadA, YadA and UspA2 have a carboxyl terminal membrane anchor made by four amphipathic β-strands and an internal α-helical region with propensity to form coiled-coils. In YadA and UspA2 these α-helices have been shown to form coiled-coils regions, which mediate oligomerisation of monomers [Hoiczyk *et al.* (2000) *EMBO J* 19:5989-5999; Cope *et al.* (1999) *J. Bacteriol.* 181:4026-4034].

The absence of cysteine residues in the mature forms of NadA is another feature shared with its homologues.

The genomic environment of NadA

The 1086bp *nadA* coding region is flanked at the 3' end by a terminator sequence while at the 5' end (Figure 12A) it shows a putative ribosome-binding site (RBS; 5'-AAGG-3') and a putative promoter region located 8 and 47 base pairs, respectively, upstream the ATG start codon.

130 bp upstream the coding region are nine repeats of the tetranucleotide TAAA (shaded black in Figure 12A), preceded by a second putative promoter with -10 and -35 regions. Because of the presence of the TAAA repeats, the gene had been listed as one of those that may undergo phase variation, even though the repeats are not in the coding region [Tettelin *et al.*]. The homologous gene *UspA2* has a tetranucleotide repeat (AGAT) located in the same position as in *nadA*, which varies in different strains [Cope *et al.* (1999) *J. Bacteriol.* 181:4026-4034].

The G+C content of the *nadA* gene and its upstream region is lower than average (45% against an average of the rest of the genome, 51.5%), suggesting acquisition of the gene by horizontal transfer.

The NadA gene and its upstream region are not present in the published sequence of the genome of serogroup A, strain Z2491 [Parkhill *et al.* (2000) *Nature* 404:502-506]. In the MenA genome, a short sequence of 16 nucleotides with no homologies in the database, replaces the nadA gene (Figure 12B), whereas the upstream and downstream genes (nmb1993 and nmb1995) are well conserved (91% and 97% identity). Analysis of the sequences immediately adjacent to the *nadA* region and absent in the Z2491 serogroup A strain shows that the segment is flanked by the TCAGAC direct repeats. This may indicate a mechanism of recombination. In the A strain the stretch of 16 nucleotides has a disrupted pair of TCAGAC repeats flanking it.

Variation in NadA genotype

Given the difference in *nadA* expression between serotypes A and B, 175 different strains of *N.meningitidis* were chosen for analysis — 150 isolates representative of the five disease-associated serogroups (A, B, C, Y and W-135) and 25 strains isolated from healthy carriers. The analysis also included one strain each of *N.gonorrhoeae, N.cinerea* and *N.lactanica.*

Bacteria were grown overnight at 37°C in a humidified atmosphere of 5% CO_2 in air on gonococcus (GC) medium agar (Difco) supplemented with Kellogg's supplement solution (0.22 M D-glucose, 0.03 M L-glutamine, 0.001 M ferric nitrate, and 0.02 M cocarboxylase) (Sigma-Aldrich Chemical Co., St. Louis, Mo.) as previously described [Knapp *et al.* (1988) *Antimicrob. Agents Chemother.* 32:765-767; Roberts *et al.* (1977) *J. Bacteriol.* 131:557-563]. One loopful of bacteria was dissolved in 500 µl of PBS and chromosomal DNA was prepared as previously described [Tinsley *et al.* (1996) *PNAS USA* 93:11109-11114].

The bacteria were screened by PCR and/or dot blot hybridization.

PCR amplification of the *nadA* genes was performed on 10 ng of chromosomal DNA using primers, mapping 350 nt upstream and downstream from the coding region (forward primer: SEQ ID 16; reverse primer: SEQ ID 17), and Platinum Hifi Taq Polymerase (GIBCO). PCR conditions were: 30 cycles of denaturation at 95°C for 30 s, annealing at 60°C for 30 s, and extension at 68°C for 1 min. PCR products were analysed on 1% agarose gel and the sizes were determined using a molecular weight marker 1Kb Plus DNA Ladder (GIBCO). The amplified fragments were purified on a Qiaquick column (Qiagen) and then automated cyclo-sequenced (Applied Biosystems model 377) by primer walking on both strands of the amplified fragment.

For dot blotting, the probe used was the whole *nadA* gene, as amplified from 2996 strain and labelled with digoxigenin using the Roche DIG High-Prime DNA Labelling and Detection Kit. 10 µl aliquot of cell suspension of each strain were boiled for 10 min. and spotted on nylon membrane (Boehringer). The membranes underwent cross-linking of DNA by 2' exposure to UV light and other standard procedures for preparation and signal detection as reported by the manufacturer.

The *nadA* gene was absent in *N.gonorrhoeae* and in the commensal species *N.lactamica* and *N.cinerea*. In *N.meningtidis*, however, 47% of isolates were positive for its presence.

PCR generated (Figure 13) a product of 1800 bp in $NadA^+$ strains MC58 (lane 1), 90/18311 (lane 2) and 2996 (lane 3). It gave a product of 400 bp in $NadA^-$ strain Z2491 and NG3/88 (lane 5). Some strains (*e.g.* 93/4286, C4678, 2022, ISS1113) gave a PCR product of 2500 bp (lane 4: L93/4286).

The presence/absence of NadA in *N.meningitidis* was correlated with strain lineage. Strains isolated from invasive meningococcal disease have been classified by multilocus enzyme electrophoresis (MLEE) into a small number of hypervirulent lineages: Electrophoretic Types ET37, ET5, cluster A4, lineage III, subgroups I, III and IV-1 [Achtman (1995) *Global epidemiology of meningococcal disease*. In *Meningococcal disease* (Cartwright, ed). John Wiley and Sons, Chichester, England. 159–175; Caugant (1998) *APMIS* 106:505-25]. Recently, a sequence-based classification, multilocus sequence typing (MLST), has been introduced, which classifies the above strains into Sequence Types ST11, ST32, ST8, ST41, ST1, ST5, ST4, respectively [Maiden *et al.* (1998) *PNAS USA* 95:3140-3145]. Strains isolated from healthy carriers fall into many different ET and ST types.

The *nadA* gene was present in 51 out of 53 strains (96%) of the hypervirulent lineages ET-5, ET-37 and cluster A4, whereas it was absent in all the tested lineage III strains. Seven of the 25 carrier strains were positive. Most of the serogroup C strains tested were positive even if not belonging to

hyper-virulent lineages. The same was true for the serogroup B strains with serotype 2a and 2b. For serogroup A, one strain belonging to subgroup III was positive whereas the other two strains belonging to subgroup IV-1 were negative.

Lineage III has only recently been introduced in Europe and USA and the geographic segregation in New Zealand for many years could have impaired its ability to acquire novel genes. For instance, mutations may have occurred in the surrounding chromosomal regions preventing Lineage III from further recombination events. Another possible explanation is that ET-5, ET-37 and Cluster A4 strains need *nadA* to achieve peak fitness whereas Lineage III isolates cannot derive any significant benefit from *nadA* insertion, thus undergoing a negative selection.

NadA is thus over-represented in three hypervirulent *N.meningitidis* lineages. It appears to be a foreign gene present in a subset of hypervirulent strains.

NadA alleles

As PCR products were differently sized (Figure 13) and most of the $NadA^+$ strains could be grouped in three different sizes, genes were sequenced for 36 strains representative of each size: 26 positive strains, 4 strains with a long PCR product, and 6 $NadA^-$ strains.

In the negative strains, a 16bp sequence was found which was identical to the sequence present in the published serogroup A genome sequence.

Analysis of the sequence of the four long PCR product strains revealed an interruption by a single copy of IS1301, interrupting the protein after 162 amino acids with a stop codon. The insertion site was identical in all four strains, but the orientation of IS1301 differed, indicating independent events. The target consensus for IS1301, 5'-AYTAG-3' was found within the NadA gene at nucleotide 472, generated by an A->G mutation, and was accompanied by a TA duplication.

In $nadA^+$ strains, gene size ranged from 1086 to 1215 bp, with consequent variation of the amino acid sequences of the encoded proteins from 362 to 405 amino acids. It was possible to cluster 22 of the 26 NadA genes into three well-defined alleles (Figures 9 & 10; Table I). The sequence of the gene within each allele is identical and overall identity between the alleles ranges from 96% to 99%. This level of conservation is surprising and suggests weak selective pressure and/or a very recent acquisition of the *nadA* gene. The latter possibility is consistent with the low G+C content of the genome in this region (see above).

Allele	Found in strains	SEQ IDs
1	MC58, BZ83, BZ169, NM066, NM119, CU385, ISS832, ISS1071, ISS1104	1,4
2	90/18311, NGP165, PMC8, M986, ISS838 and 961-5945	2,5
3	C11, 973-1720, ISS759, F6124, 2996, 8047, NMB	3,6

The sequences shown in Figure 9A assume that the N-terminus amino acid is the first Met in the open reading frame (SEQ IDs 4 to 6), but the second Met (residue 3 in SEQ IDs 4 to 6) has a better-positioned Shine-Dalgarno motif (Figure 9B). Sequences starting from the second Met codon are thus preferred (SEQ IDs 1 to 3).

Allele 1 codes for a protein of 362 amino acids (SEQ ID 1) and includes strain MC58 and all the ET-5 positive strains sequenced. The other five strains belonging to allele 1 were very recent isolates and they have not been ET-typed yet, although serotype and serosubtype classification (B:15:P1.7 and B:4:P1.15) of these strains suggests affiliation of these strains to the ET-5 complex.

Allele 2 codes for a protein of 398 amino acids (SEQ ID 2) resulting from the addition of 2 aa after residue 268 (numbering according to SEQ ID 1), addition of 41 aa after residue 271, and deletion of 7 aa after residue 122, resulting in the deletion of the first heptad repeat of the leucine zipper domain. Leucine residues at a fixed spacing of seven residues commonly identify leucine zippers. One leucine in the repeats has frequently been replaced mostly by Met, Val or Ile. In this case allele 2 could use the Ile upstream or downstream to form the leucine zipper motif.

Allele 3 codes for a protein of 405 amino acids (SEQ ID 3) and, like allele 2, contains 43 extra amino acids at residues 268 and 271 but differs from allele 2 by not having the 7aa deletion after residue 122. Allele 3 is found in serogroup A, B and C strains.

The remaining 4/26 positive strains (ISS1024, ISS759, 973-1720, 95330; marked with * in Table 1) contain minor variants of alleles 1 to 3:

- Serogroup C strain ISS1024 has a variant of allele 2 with a single heptad repeat deletion at residues 229-235 (SEQ IDs 7/8). This sequence was originally classified as a fourth allele but has been re-classified as a variant of allele 2. Allele 2 is thus found in all ET-37 strains, one strain of cluster A4 and three additional non-ET-typed serogroup C strains.
- Serogroup C strains ISS759 and 973-1720 both contain a variant of allele 3 with a single amino acid mutation in the leader peptide (SEQ IDs 9/10) resulting from a single nucleotide mutation. Among all allele 3 strains, only 973-1720 belongs to a hypervirulent strain (cluster A4).
- Serogroup B strain 95330 contains a recombinant (chimera) of alleles 1 and 2 (SEQ IDs 11/12), with *nadA* being a fusion between the N-terminal portion of allele 2 and the C-terminal segment of allele 1. The putative site of recombination is located approximately between residues 141 and 265 of the protein.

All insertions and deletions happen in the coiled-coil region and involve 7 or 41 amino acids which, representing 2 or 6 turns of the α-helix, allows for variations in length of the coiled coil region without disturbing the overall structure. Furthermore, the deletion in ISS1024 results in the loss of the first heptad repeat of the leucine zipper domain but does not destroy the domain because leucine residues at a fixed spacing of seven residues can be replaced mostly by Met, Val or Ile. In this case allele 2 could use the Ile upstream or downstream to form the leucine zipper motif (Figure 11).

Any of these various NadA sequences and alleles can be used in accordance with the invention.

When sequence analysis was extended to the putative promoter and terminator regions (50bp upstream, 350bp downstream), variations were found only in the in the 5' region. Three Italian strains (ISS1071, ISS832 and ISS1104) differed for a single base mutation while in strain 961-5945 there was a 7 base differences (indicated with * in Figure 10). Variations were also found in the 5' regions

where the TAAA tetranucleotide was repeated from 4 to 12 times in different strains (Table 1). The number of repeats was variable also within each allele (Table 1).

Further work was performed on carrier strains isolated from healthy individuals by oro-pharyngeal swab. Some strains, even if described as carriers, belong to hypervirulent clusters, and NadA was found in all such carrier strains as described above (*i.e.* allele 1 in the ET-5 strains and allele 2 in the ET-37 strains).

NadA was also found in five carrier strains (NGE28, 65/96, 149/96, 16269, 16282) which do not belong to a hypervirulent cluster. These five strains shared a sequence (SEQ IDs 13 & 14) which was not found in strains isolated from patients. This allele is referred to as 'allele C' (carrier).

An alignment of allele C with alleles 1 to 3 is shown in Figure 9C. Disruption in the coiled-coil segments of the protein is evident.

Unlike alleles 1 to 3, allele C protein does not readily form a high molecular aggregate when expressed in *E.coli*. Like alleles 1 to 3, however, allele C is exposed on the surface of *N.meningitidis*, because it is a target for bactericidal antibody raised against itself. However, these antibodies are not bactericidal against strains carrying alleles 1 to 3; similarly, antibodies raised against alleles 1 to 3 are not bactericidal against allele C strains.

NadA oligomers on the cell surface

WO01/64922 reports that NadA forms oligomeric structures. To study NadA oligomers in more detail, whole cell lysates of *N.meningitidis* were probed by Western blot.

Bacterial colonies [strains MC58 (allele 1), 90/18311 (allele 2), 2996 (allele 3), L93/4286 (IS1301 insertion) and NG3/88 (*nadA*$^-$)] were grown to stationary phase in GC broth supplemented with 0.3% glucose. Samples were taken at different times, pelleted by centrifugation at 3000 x g for 10 min, and resuspended in PBS and thawed/frozen up to bacterial lysis. Equal amounts of proteins were subjected to SDS-PAGE on 12.5% polyacrylamide gels and electrotransferred onto nitrocellulose membranes.

To prepare anti-NadA polyclonal serum, recombinant NadA was expressed and purified. Sequences encoding the three *nadA* alleles (allele 1: aa 24-362; allele 2: aa 24-343; allele 3: aa 24-350), were amplified by PCR on chromosomal DNA and cloned into pET21b+ vector (Novagen). The plasmids were transformed in *E.coli* BL21 (DE3) to express the proteins as C-terminal histidine fusions. Protein expression was induced at 30°C by adding 1mM IPTG at OD_{600nm} 0.3 and growing the bacteria for an additional 3 h; expression was evaluated by SDS-PAGE. Recombinant fusion proteins were purified by affinity chromatography on Ni^{2+}-conjugated chelating fast-flow Sepharose 4B resin. 20 μg of purified protein was used to immunise six-week-old CD1 female mice (4 to 6 per group). Proteins were given intraperitoneally, with complete Freund's adjuvant (CFA) for the first dose and incomplete Freund's adjuvant (IFA) for the second (day 21) and third (day 35) booster doses. Bleed out samples were taken on day 49 and used for the serological analysis.

The blots showed a high molecular weight reactive band in strains MC58 (Figure 14, lane 1), 90/18311 (lane 2) and 2996 (lane 3). The band was absent in strain NG3/88 (lane 5). Boiling of the sample buffer up to 40 minutes did not change the pattern. The different size of the proteins was consistent with the size of the alleles. Given the expected size ranging from 35 to 40 kDa of monomeric proteins, the high MW of the observed band could be explained by the presence of an oligomeric form of NadA. This possibility is supported by the fact that in a strain containing the IS1301 insertion, coding for a shorter protein of 162 amino acids and lacking most of the coiled-coil region, the high MW reactive band was absent and replaced by a band of 14.5 kDa (Figure 14, lane 4), consistent with the predicted molecular weight of the processed monomeric protein.

Although the oligomeric protein was found in all strains containing a functional gene, expression levels varied from strain to strain (Table I). Moreover, the amount of NadA protein varied within the same strain during growth.

Four different strains (MC58, 2996, C11, F6124), chosen as representative of diverse overall NadA expression level, were followed during growth up to stationary phase. Figure 15 shows growth of two of the tested strains (15A: MC58, with low NadA expression; 15B: 2996, with high NadA expression), with the curve showing OD_{600}. Western blots of samples taken at each point of the OD_{600} growth curve showed that the NadA band was barely visible at the beginning of the growth and became more intense during growth, up to its maximum, at stationary phase. All strains analysed showed the same growth-phase dependent behaviour.

High MW NadA was also seen in western blots of outer membrane vesicles, consistent with NadA being anchored to the outer membrane.

Similarly, FACS analysis on live bacteria during log-phase growth showed that NadA was available for antibody binding on the surface of the bacteria. FACS intensity in a strain with a poylsaccharide capsule (strain NMB) was reduced 1 log in comparison to an isogenic non-encapsulated mutant strain (M7), but the protein was surface-exposed and available for binding in both strains (Figure 16).

NadA forms surface-exposed oligomers, which are stable to heat, SDS and reduction with β-mercaptoethanol. As the mature form of the lacks cysteine residues, disulphide bond formation cannot be involved in this phenomenon; rather this is consistent with the predicted coiled-coil structure and the presence of leucine zipper motifs that might mediate intermolecular interactions between monomers [Lupas (1996) *Trends Biochem. Sci.* 21:375-382; O'Shea *et al.* (1991) *Science* 254:539-544]. The size of the oligomers is approximately 170 kDa, suggesting a tetrameric structure [WO01/64922]. However, a rigid coiled-coil structure is likely to have an anomalous migration is SDS PAGE and therefore the 170kDa form may be a trimer.

Protective immunogenicity

Polyclonal anti-NadA serum was tested for bactericidal activity as previously described [Pizza *et al.* (2000); Peeters *et al.* (1999) *Vaccine* 17:2702-2712], with pooled baby rabbit serum (CedarLane) used as complement source. Serum bactericidal titer, was defined as the serum dilution resulting in a 50% decrease in colony forming units (CFU) per ml after 60 minutes incubation of bacteria in the

reaction mixture, compared to control CFU per ml at time 0. Typically, bacteria incubated with the negative control antibody in the presence of complement showed a 150 to 200% increase in CFU/ml during the 60 min. of incubation.

Results were as follows:

Strain	NadA expression	Allele	Bactericidal titre
2996	+++	3	32768
C11	+++	3	16384
F6124	+	3	4096
MC58	+	1	8192
BZ232	-	-	<4
NGH38	-	-	<4

As shown, the serum induced complement-mediated killing of all strains that have the *nadA* gene, and was inactive against the strains that do not have the gene. However, bactericidal titres varied between strains. Titres were higher against strains expressing higher amounts of protein. This result was confirmed when titres were determined in the early and late phase of growth (Figure 15).

To check whether the differences in the bactericidal activity were due to different allele sequences, immune sera, raised against the three NadA types, were produced and used in a cross bactericidal assay. The results obtained with the antisera were similar to those shown above, suggesting that the bactericidal activity is not influenced by the allele diversity but rather to the antigen expression level.

The ability of immune sera to protect animals from bacteremia during infection was also tested, using the infant rat model. The sera used were obtained by immunising guinea pigs with 50µg purified rNadA (allele 3). Immunisation of outbred Wistar rats (5 to 7 days old) was performed subcutaneously together CFA for the first dose and IFA for the further three doses (days 28, 56, 84). Bleed out samples were taken on day 105 and used for the animal protection assay.

Two experiments were performed using two different MenB strains (8047 and 2996). Each strain has been serially passaged three times in infant rats. In experiment 1, groups of four rats were challenged intraperitoneally with 100µl of a mix of (a) bacteria from strain 8047 ($7x10^3$ CFU per rat) and (b) heat inactivated guinea pig antiserum or anti-capsule control mAb (SEAM 3 [Van Der Ley *et al.* (1992) *Infect. Immun.* 60:3156]). In experiment 2, group of six rats were treated with the control mAb or with different dilutions of guinea pig antiserum at time 0. Two hours later, they were challenged with the 2996 bacteria ($5.6x10^3$ CFU per rat). In both experiments, blood cultures were obtained 18 h after the challenge by puncturing the heart with a syringe and needle containing approximately 25 U of heparin without preservative. Bacteria numbers in the blood cultures were obtained by plating out 1, 10, and 100 µl of blood onto chocolate agar overnight. For calculation of geometric mean CFU/ml, animals with sterile cultures were assigned a value of 1 CFU/ml.

Results were as follows:

Exp[t]	Treatment	Blood culture at 18 hours	
		Positive/Total	CFU/ml (10^3)
1	Anti-capsular mAb (2μg/rat)	0/4	<0.001
	Anti-NadA antiserum (1:5 dilution)	0/4	<0.001
	PBS + 1% BSA	5/5	40.17
2	Anti-capsular mAb (20μg/rat)	1/6	0.003
	Anti-NadA antiserum (1:5 dilution)	1/6	0.002
	Anti-NadA antiserum (1:25 dilution)	3/6	0.035
	Pre-immune NadA serum	6/6	1.683

Thus anti-NadA antiserum is highly protective in this assay.

Overall, therefore, NadA has several attributes of being a good vaccine antigen: (i) it is a surface-exposed molecule, potentially involved in bacterial adhesion; (ii) it is present in at least 50% of the disease-associated strains and in almost 100% of three hypervirulent lineages; (iii) it elicits protective and bactericidal antibodies in laboratory animals; and (iv) each allele induces cross-bactericidal antibodies.

ORF40

ORF40 shows homology to Hsf and its allelic variant Hia, both adhesins of *Haemophilus influenzae*. The different size among Hia, Hsf and ORF40 is in part explained by the presence of three copies of a large repeated domain in Hsf, which is present in single copy in Hia and only partially in ORF40 (Figure 7). In MenB, ORF40 is found on the outer membrane as a protein of about 200 kDa (*cf.* predicted MW of 59 kDa for mature protein).

App

App shows homology (Figure 8) to the adhesion and penetration protein Hap of *H.influenzae*, which is an adhesin with a serine-protease activity that undergoes autoproteolytic cleavage and extracellular release [Hendrixson *et al.* (1997) *Mol Microbiol* 26:505-518]. Uncleaved surface-associated Hap mediates adherence to epithelial cells and promotes bacterial aggregation and colonisation.

In *N.meningitidis*, App is exported to the outer membrane, processed and secreted. Both Hap and App belong to the autotransporter family which comprises proteins from gram-negative bacteria characterized by a distinct mechanism of secretion. This system was first described for IgA1 protease of *N.gonorrhoeae*, which is considered the prototype of this family. Proteins of the autotransporter family have been implicated in the virulence of many gram-negative pathogens [Henderson & Nataro (2001) *Infect Immun* 69:1231-1243]. They are synthesized as large precursor proteins comprising at least three functional domains: a typical N-terminal leader sequence, an internal domain (passenger domain) and a C-terminal domain (translocator domain or β-domain). The leader sequence mediates the export (sec-dependent) of the protein to the periplasm. Subsequently the translocator domain inserts into the outer membrane forming a β-barrel pore to allow the export of the passenger domain. Once at the bacterial surface, the passenger domain can be cleaved and released into the environment. Cleavage can occur by an autoproteolytic event directed by protease activity in the

passenger domain itself. Passenger domains of autotransporters are widely divergent, reflecting their remarkably disparate roles. On the contrary the β-domains display high degree of conservation consistent with their conserved function.

App possesses the prevailing domains of the autotransporter proteins as well as the conserved serine protease motif (GDSGSP). It has been shown that this motif is responsible for cleavage of human IgA by the Neisseria IgA1 proteases and for autoproteolytic cleavage of Hap protein of *H.influenzae*. App has been shown to be a conserved antigen among meningococci, to be expressed during infection and carriage, to stimulate B cells and T cells, and to induces a bactericidal antibody response [Hadi *et al.* (2001) *Mol. Microbiol.* 41:611-623; Van Ulsen *et al.* (2001) *FEMS Immunol Med Microbiol* 32:53-64].

In serogroup B strain 2996, App has 1454 amino acids and a predicted MW of 159,965 Da. Figure 26 shows the protein's predicted structural features. Three domains can be seen: domain 1 (amino acids 1-42) is the signal peptide; domain 2 is the passenger domain, which is the functionally active protein; domain 3 is the C-terminal translocator domain with β barrel structure.

At the N-terminus of the passenger domain, His-115, Asp-158 and Ser-267 correspond to the serine protease catalytic triad His-98, Asp-140 and Ser-243 from Hap [Fink *et al.* (2001) *J Biol Chem* 276:39492-39500]. Residues 285-302 are a putative ATP/GTP-binding site (P loop), which suggests a mechanism of energy coupling for outer membrane translocation. Towards the C-terminus of the passenger domain, two Arg-rich regions are present. The first (RRSRR) is residues 934-938 and the second (RRARR) begins at residue 1149. These motifs are reminiscent of known targets for trypsin-like proteolytic cleavage sites such as the one in diphtheria toxin and those upstream of the auto-cleavage sites of *H.influenzae* Hap, *N.gonorrhoeae* IgA-protease and *B.pertussis* FhaB (Figure 26, box 1). Downstream of the Arg-rich regions are motifs ^{954}NTL^{956} and ^{1176}NSG^{1178}, which are identical or similar to the cleavage sites in autotransporters Ssp (*Serratia marcescens*), Prn (*Bordetella bronchiseptica*), Brka (*Bordetella pertussis*) [Jose *et al.* (1995) *Mol. Microbiol.* 18:378-380] and Hap (*H.influenzae*) (Figure 26, box 2). Together, these sequence motifs suggest that the two motifs ^{954}NTL^{956} and ^{1176}NSG^{1178} and the RR(A,S,R)$_2$RR pattern could act as signals for correct localisation of downstream processing sites.

Further analysis of the App sequence shows a proline-rich region, where the dipeptide motif PQ is repeated four times beginning at residue 1156. A search for homology to known protein sequences reveals some similarity to the surface proteins of *S.pneumonie* PspA and PspC and to a proline-rich region of the *B.pertussis* outer membrane protein p69 pertactin, where the (PQP)$_5$ motif is located in a loop containing the major immunoprotective epitope.

Finally, the last three amino acids of App (YRW) are identical to those of Hap where they have been described as crucial for outer membrane localisation and protein stability [Hendrixson *et al.*, 1997].

Expression in* E.coli *without fusion partners

ORF40, App and NadA full-length genes were cloned in pET21b+ vector and the plasmids were transformed in *E.coli* BL21(DE3) in order to express the genes under control of T7 promoter.

Expression was achieved activating the promoter with IPTG or under non-induced conditions. Localisation and surface-exposure of the proteins were assayed by cell-fractionation experiments (SDS-PAGE and Western blot), FACS analysis and whole-cell immunoblot. As shown in figures 1 to 3, all the three proteins are translocated to the surface of *E.coli*:

- ORF40 is expressed as monomeric form and possibly forms also multimers (Figure 1).
- App is exported to *E.coli* outer membrane as a precursor of about 160 kDa, that is processed and secreted in the culture supernatant (Figure 2).
- NadA is found to the be present in the outer membrane fraction as a single high molecular weight band of approximately 180 kDa. This probably corresponds to an oligomeric form of the protein. Such a band is absent in *E. coli* expressing intracellular NadA (Figure 3).

App expression was studied in more detail.

N.meningitidis strain 2996 genomic DNA was prepared as previously described [Tinsley & Nassif (1996) *PNAS USA* 93:11109-11114]. DNA devoid of the sequence coding for the signal peptide (amino acids 1 to 42) and of the STOP codon was amplified using PCR primers SEQ IDs 18 & 19 followed by digestion with *Nhe*I and *Xho*I and insertion into the *NheI/XhoI* sites of the pET-21b expression vector, to give 'pET-App-His' (Figure 27). This plasmid was introduced into *E.coli* BL21(DE3) and used for the expression of a C-terminal His-tagged fusion protein which was purified and used to raise antibodies. The full-length *app* gene was amplified and cloned in a similar way, using PCR primers SEQ IDs 20 & 21, to give plasmid 'pET-App'.

Plasmids were introduced into *E.coli* BL21(DE3) and expression induced by addition of 1mM IPTG. The expressed protein was detected by western blotting (Figure 28, lane 1). To verify that the protein was exported to the *E.coli* surface, FACS (Figure 29) and immunofluorescence microscopy (Figure 30) were used. The FACS analysis showed positive surface expression on the pET-App transformants (full-length gene) but no surface expression with App-His (no signal peptide) or with the empty vector. The immunofluorescence results agreed with FACS. Therefore expression of the full-length *app* gene resulted in the export of App to the surface of *E.coli*, but deletion of the first 42 amino acids abolished surface-localisation.

Western blot analysis of outer membrane proteins from pET-App transformants revealed a specific reactive band of ~160 kDa (Figure 28, lane 1), corresponding to the predicted molecular weight of the full-length protein. A corresponding band was missing in the outer membrane fraction from untransformed controls (lane 3). Western blot analysis of culture supernatants revealed a secreted protein of ~100 kDa with pET-App (lane 2) that was absent with the untransformed controls (lane 4). Sometimes a very weak band was also detected at ~140 kDa in pET-App transformants.

Therefore the full length *app* gene when introduced into *E.coli* induces expression of an App protein which is exported to the outer membrane, cleaved and released into the culture supernatant.

Native expression can influence the quality of the immune response

To evaluate the role of protein conformation on induction of an immune response, outer membrane vesicles from *E.coli* expressing ORF40, App or NadA were isolated and used to immunise mice. Sera were tested for bactericidal activity and results compared with those obtained with the fusion proteins. The bactericidal response (strain 2996) was improved 5-10 fold when the proteins are produced in their "native" form in OMVs:

Antigen	Bactericidal titres *	
	Fusion protein	*E.coli* OMV
ORF40	256	2048
App	64	1024
NadA	32768	>65536

* Titres expressed as the reciprocal of the serum dilution yielding ~50% bacteria killing

App autoproteolytic cleavage

E.coli pET-App transformants secrete a 100kDa product into culture supernatant and show a 160kDa surface product. To test whether the secreted App product derives from an autoproteolytic process, one of the putative catalytic residues (Ser-267) was replaced with Ala.

The pET-AppS267A mutant was obtained by site-directed mutagenesis using the QuikChange kit (Stratagene) and primers SEQ IDs 22 & 23.

SDS-PAGE analysis of total proteins from pET-AppS267A transformants (figure 31, lane 2) showed a protein similar in size to pET-App transformants (lane 1). The protein was shown to be surface exposed by FACS analysis (Figure 29). Western blot analysis of culture supernatants showed App in pET-App transformants (Figure 32, lane 1) but not in pET-AppS267A transformants (lane 2).

Mutation of Ser-267 to Ala thus abolishes processing and secretion of the App precursor, which remains cell-associated. These data suggest that App has a serine protease activity that is responsible for autoproteolytic processing and release in the supernatant of the secreted App domain.

Cleavage at ^{954}NTL^{956} would leave a fragment with predicted molecular weight of 104190 Da. Cleavage at ^{1176}NSG^{1178} would give a 128798 Da fragment. These two predicted fragments may match the two bands of ~140 and ~100 kDa observed in culture supernatants. Cleavage may occur first to give the ~140 kDa fragment and then second to give the 100 kDa fragment. The β domain of App would thus begin at residue 1177.

NadA, ORF40 and App function as adhesins

Example 22 of international patent application WO01/64922 discloses that NadA expression in *E.coli* makes the transformed bacterium adhere to human epithelial cells. The adherent phenotype has been further studied for NadA and also for App and ORF40.

E.coli BL21(DE3) bacteria (10^8 CFU), grown under non-induced or induced conditions, were inoculated onto Chang human epithelial monolayers (10^5 cells) and incubated at 37°C for 1 or 2

hours. Cells were then incubated with rabbit anti-*E.coli* and PE-conjugate secondary antibody. Adhesion was detected by FACS as specific fluorescence intensity associated to Chang cells. Positive controls were *E.coli* DH5 expressing *hsf* (DH5/pDC601)); negative controls were BL21(DE3)/pET21b and DH5a/pT7-7. The results in figure 4 show that the ability of the recombinant *E.coli* strains to adhere to cultured epithelial cells is associated with expression of these three proteins.

To confirm that these three proteins are able to promote interaction with host cells, the recombinant proteins themselves were investigated for binding to epithelial cells. 10^5 Chang human epithelial cells (Wong-Kilbourne derivative, clone 1-5c-4, human conjunctiva) were incubated at 4°C for 30 minutes with medium alone or with different concentration of ORF40 (150μg/ml), App (150μg/ml) or NadA (300μg/ml), or with GNA2132 (300μg/ml) as negative control [see Pizza *et al.* (2000)]. Binding was detected by FACS using polyclonal antisera against the single recombinant proteins and a secondary PE-conjugate antibody. The FACS signal shifts (Figure 5) show that the three proteins are able to bind to human epithelial cells, whereas purified GNA2132 (negative control) does not.

Figure 6A shows that binding increases in a dose-dependent manner. Binding of NadA reaches a plateau at around 200μg/ml. GNA2132 fails to bind even at 400μg/ml (Figure 6B). Data in Figure 6 are mean fluorescent intensity (MFI) values plotted against protein concentration (μg/ml).

Using FACS, binding of NadA to cells was also seen with Hep-2 and MOLT-4 cells, but not with HeLa, A549, Hec-1B, Hep-G2, CHO or HUVEC cells. Adhesion to Chang cells could be abolished by treating the cells with pronase, indicating that the human receptor for NadA is a protein.

Adhesion of purified NadA protein to Chang conjunctiva cells was also observed using immunofluorescence microscopy. The protein (lacking its C-terminal anchor domain) was incubated with Chang cells at 37°C in complete culture medium for 3 hours at various concentrations. Cells were then washed, fixed, and analysed by laser confocal microscopy after staining with anti-NadA mouse polyclonal antibodies and secondary Texas-red coupled anti-mouse IgG antibodies. No binding was seen at 0nM (figure 17A), but binding was evident at 170nM (17B) and 280nM (17C), with clustering evident at higher concentrations. In contrast, no binding of NadA was seen with HeLa cells, even at 280nM protein (17D).

Binding was much more evident at 37°C (figure 18A) than at 4°C (figure 18B). The dot-like structures seen at 4°C, compared to clusters at 37°C, suggest that lateral interactions between NadA monomers are temperature-dependent (influenced by membrane fluidity).

To distinguish surface and endocytosed protein, saponin detergent was added during the staining procedure. Intracellular clusters having the size of endosomes were more evident (arrow) when saponin was used, but a high proportion of protein remained on the cell surface (figure 19).

Immunofluorescence also revealed that NadA binds to monocytes (figure 20A). NadA alone (no staining antibody; 20B) and NadA stained with pre-immune serum (20C) were not visible. At high magnification, evidence of uptake into vesicles (either endosomes or phagosomes) was seen.

Figure 21 shows that murine macrophages (raw 264.7) bind and endocytose NadA (125nM, 3 hours, 37°C; cells cultured in DMEM).

Heating NadA at 95°C for 15 minutes prior to incubation removed its ability to bind to monocytes, as measured by secretion of IL-α by the cells (figure 22). The stimulatory activity of NadA preparations is thus heat-labile. Stimulatory activity was also blocked by the use of anti-CD14 (figure 23) or by the removal of NadA from the preparations using bead-immobilised anti-NadA.

Immunofluorescence microscopy was also used to detect binding of *E.coli* expressing NadA. Transformed *E.coli* bound strongly (figure 24A) whereas untransformed bacteria did not (24B). IL-α release by monocytes was over 1.5x higher using the transformed *E.coli* than the untransformed bacteria at a bacteria/monocyte ratio of 40:1.

Transformed *E.coli* were bound to glass cover slips, fixed and double-stained with anti-NadA (figure 25A) and anti-*E.coli* antibodies (25B). When both were used, patches of anti-NadA were visible, suggesting that NadA tends to form aggregates on the bacterial surface, which hamper the interaction of antibodies with other surface antigens.

Looking at App, recombinant *E.coli* strains were incubated with monolayers of Chang conjunctiva epithelial cells (Wong-Kilbourne derivative, clone 1-5c-4 [human conjunctiva], ATCC CCL 20.2) and adhesion was analysed using FACS. Cells obtained from confluent monolayers were seeded at 10^5 cells per well in 12-well tissue culture plates and incubated for 24 hours. Cultures of bacteria after IPTG induction were washed twice in PBS and resuspended in DMEM+1% FBS to a concentration of $5x10^8$ bacteria per ml. Aliquots of 1 ml of each strain were added to monolayer cultures of Chang cells and incubated for 3 hours at 37°C in 5% CO_2. Non-adherent bacteria were removed by washing three times with PBS, and 300 µl of cell dissociation solution (Sigma) were added to each microtitre well. Incubation was continued at 37°C for 10 minutes. Cells were harvested and then incubated for 1 hour at 4°C with rabbit polyclonal anti-*E.coli* antiserum (DAKO). Cells were washed twice in PBS+5% FBS and incubated for 30 mintues at 4°C with R-Phycoerythrin-conjugated anti-rabbit IgG (Jackson ImmunoResearch Laboratories). Cells were then washed in PBS+5% FBS and resuspended in 100µl PBS. Fluorescence was measured with FACSCalibur flow cytometer (Becton Dickinson). For each of fluorescence profile, 10000 cells were analysed.

The results reported in Figure 33 show pET-App transformants were able to adhere to Chang cells, giving a fluorescence shift of 90.3%. S267A transformants were also able to adhere (91.0%). Untransformed *E.coli* were unable to adhere to Chang cells (bottom FACS plot).

As for NadA, FACS results were in agreement with immunofluorescence microscopy data. As shown in Figures 34A & 34B, pET-App transformants incubated with monolayers demonstrated high levels of adhesion to epithelial cells and visible bacteria-bacteria aggregation. For the S267A mutant, adhesion and bacterial aggregation were increased (34C & 34D). Untransformed controls showed no adhesion (34G). Deletion of the first 42 amino acids also abolished adhesion.

In contrast to Chang epithelial cells, no adhesion was seen when HUVEC endothelial cells were tested with pET-App transformants. To cause sepsis and meningitis, *N.meningitidis* has to interact

with human endothelial cells. App may thus be involved in the first step of colonisation at the level of human respiratory epithelial mucosa, rather than in pathological endothelial colonisation.

Localization and specificity of App binding activity.

To identify the binding region of App, a chimeric protein named Appβ was used. This protein consists of the C-terminal domain of App (amino acids 1077 to 1454) fused to the leader peptide of IgA1 protease of *N.gonorrhoeae*. The gonococcal leader sequence was chosen because it has been well characterized and is functional in *E.coli*. Plasmid pET-Appβ contains a 1.1 kbp DNA fragment amplified by PCR using SEQ IDs 26 & 27.

The pET-Appβ construct was introduced into *E.coli* BL21(DE3). FACS localisation studies confirmed that Appβ was localized on the *E.coli* surface. The *in vitro* adhesion assay using Chang epithelial cells showed adhesion by immunofluorescence (Figure 34E & 34F). FACS analysis showed that the pET-Appβ transformants were still able to adhere to epithelial cells but at lower levels (74.2% shift) than pET-App transformants.

These results indicate that the App binding domain is located in its C-terminal region, in the 100mer fragment between residues 1077 and 1176.

Purified recombinant proteins were also studied. App-α-His consists of the N-terminal portion of App (amino acids 43-1084) fused to a poly-His tag. Plasmid pET-Appα-His contains a *NheI/XhoI* 3.1 kbp fragment amplified by PCR with SEQ IDs 24 & 25. The binding activity of the purified recombinant App-α-His was compared to that of App-His by FACS binding assays. Chang cells were incubated with increased concentrations of recombinant App proteins or lipoprotein NMB2132-His (negative control). Binding of App-His (♦) increased in a dose-dependent manner and reached a plateau at a concentration of ~50μg/ml whereas the binding of Appα-his (■) was very low (Figure 35). The control NMB2132-His (▲) failed to bind Chang cells.

To explore the biochemical nature of the molecule involved in interaction with App, the Chang cells were treated with pronase or phospholipase A2 before the binding experiments. 10^5 cells per well were placed in microplates and incubated in FCS-free DMEM at 37°C in 5% CO_2 for 30 minutes with (a) pronase at 250, 500, or 1000 μg/ml or (b) phospholipase A2 at 50, 200, or 800 μg/ml. After enzymatic incubation, an equal volume of complete medium was added to each well to stop the reaction. Cells were subsequently mixed with 100 μg/ml App-His or medium alone and incubated for 1 hours at 4°C. As shown in Figure 36, pronase treatment (left-hand columns) markedly reduced the binding of App-His protein to Chang cells, while treatment with phospolipase A2 (right-hand columns) did not reduce the binding. The receptor for App on Chang cells is thus proteinaceous.

Adhesion to different cell lines were also tested (Figure 37). After incubation of cultured cells with three different concentrations of App-His (100, 25 & 6.25 μg/ml) high level binding to Chang cells and HepG2 cells was seen, a moderate level of binding to A-549 cells, and minimal binding to HeLa cells. No binding was observed to Hec-1-B, Hep-2, 16HBE14o epithelial cell lines or to HUVEC endothelial cells.

App knockout

After the work on *E.coli* suggesting an adhesin role for App, an isogenic mutant strain of *N.meningitidis* was constructed. The starting strain was MC58. Its *app* gene was truncated and replaced with an antibiotic cassette by transforming the parent strain with the plasmid pBSUDAppERM, which contains a truncated app gene and the ermC gene (erythromycin resistance) for allelic exchange. Briefly, 600bp of the upstream flanking region including the start codon and 700 bp downstream flanking region including the stop codon were amplified from MC58 using primers SEQ IDs 28 to 31. Fragments were cloned into pBluescript and transformed into *E.coli* DH5 using standard techniques. Once all subcloning was complete, naturally competent *N.meningitidis* strain MC58 was transformed by selecting a few colonies grown overnight on GC agar plates and mixing them with 20 µl of 10mM TrisHCl pH8.5 containing 1 µg of plasmid DNA. The mixture was spotted onto a GC agar plate, incubated for 6 hrs at 37°C, 5% CO2 then diluted in PBS and spread on GC agar plates containing 5 µg/ml erythromycin. The deletion app gene in the genome of MC58 was confirmed by PCR. Lack of App expression was confirmed by Western blot analysis.

Adhesion of wildtype MC58 and the isogenic MC58Δapp mutant strain was evaluated on Chang cells. There was a ~10 fold reduction (ranging from 3- to 27-fold in different experiments) of the association of the knockout mutant compared with the wild type strain (Figure 38). No difference was observed between the *app*⁻ mutant and the parental strain with Hep2 and 16HBE14o cell lines and with HUVEC endothelial cells, confirming that App does not mediate adhesion to these cells.

No non-pilus adhesins which contribute to adhesion of *N.meningitidis* in a capsulated background have previously been reported.

App expression was studied in *N.meningitidis* MC58. Colonies from plates grown overnight were diluted in GC broth and incubated at 37°C with 5% CO_2. Samples were taken when OD_{620nm} = 0.5 (mid log phase) and 0.8 (stationary phase) and analysed by western blot. Two bands with apparent molecular weights ~160 and ~140 kDa were detected in whole cells lysates of log phase bacteria (Figure 39, lane 1), while stationary phase bacteria showed only a faint band at ~140 kDa (lane 3). As expected, no App was observed in the ΔApp mutant (lanes 2 & 4).

In marked contrast, supernatant samples of wild-type MC58 showed a band at ~140 kDa and its amount was higher in stationary phase than in log phase (Figure 40, lanes 3 & 1). The stationary phase sample also showed a reactive band at ~100 kDa.

It will be understood that the invention is described above by way of example only and modifications may be made whilst remaining within the scope and spirit of the invention.

TABLE I — Characteristics of 26 *N.meningitidis* strains and their *nadA* gene allele

Strain	Serogroup type:subtype	Clonal group	nadA allele	(TAAA) repeats	NadA expression
64/69	NG:15:P1.7,16	ET-5	1	4	+
BZ83	B:15	ET-5	1	5	+++
CU385	B:4:P1.15	ET-5	1	6	++
MC58	B:15:P1.7,16b	ET-5	1	9	+
BZ169	B:15:P1.16	ET-5	1	12	++
95330*	B:4:P1.15	ET-5	1	9	nd
ISS1104	B:15:P1.7,16	nd	1	4	+
ISS1071	B:15:P1.7,16	nd	1	5	+++
ISS832	B:15:P1.7	nd	1	5	++
NM119	B.4.P1.15	nd	1	6	nd
NM066	B:15:P1.7,16	nd	1	12	nd
90/18311	C:NT:P1.5	ET-37	2	9	++
NGP165	B:NT:P1.2	ET-37	2	9	++
FAM18	C:2a:P1.5,2	ET-37	2	9	nd
M986	B:2a:P1.5,2	ET-37	2	12	++
ISS1024*	C:2b:P1.5	nd	2	9	++
ISS838	C:2a:P1.5,2	nd	2	6	++
PMC8	C:	nd	2	10	++++
961-5945	B:2b:P1.21,16	A4	2	12	+++
ISS759*	C:2b:P1.2	nd	3	8	++++
F6124	A	Subgroup III	3	9	+
NMB	B:2b:P1.5,2	nd	3	12	++
8047	B:2b:P1.2	nd	3	12	+++
2996	B:2b:P1.5-1,2	nd	3	12	+++
C11	C:NT:P1.1	nd	3	12	+++
973-1720*	C:2b:P1.2	A4	3	12	+++

* indicates that the strain carriers a minor variant of the relevant allele

nd = not done

TABLE II — Characteristics of *N.meningitidis* strains analysed for NadA expression

ST	ET	Strain	Year	Serogroup:type:subtype	Country	Disease	*NadA* gene
74	ET5	MC58	1985	B:15:P1.7,16b	UK	case	+
32	ET5	H44/76	1976	B:15:P1.7,16	Norway	case	+
32	ET5	BZ169	1985	B:15:P1.16	Netherlands	case	+
32	ET5	30/00	2000	B:15:P1.7,16	Norway	case	+
33	ET5	N44/89	1989	B:4,7:P1.19,15	Brazil	case	+
34	ET5	BZ83	1984	B:15	Netherlands	case	+
-	ET5	72/00	2000	B:15:P1.7,13	Norway	case	+
-	ET5	39/99	1999	C:15:P1.7,16	Norway	case	+
-	ET5	M4102	1996	B:ND	USA	case	+
-	ET5	95330	1995	B:4:P1.15	Canada	case	+
-	ET5	2201731	1993	NG:4:P1.15	Iceland	carrier	+
-	ET5	64/96	1996	NG:15:P1.7,16	Norway	carrier	+
-	ET5	CU385	1980	B:4:P1.15	Cuba	case	+
-	ET5	8680	1987	B	Chile	case	+
-	ET5	204/92	1992	B	Cuba	case	+
-	ET5	EG329	1985	B	Germany	case	+
-	ET5	NG080	1981	B	Norway	case	+
-	ET5	NG144/82	1982	B	Norway	case	+
-	ET5	NG PB24	1985	B	Norway	case	+
-	ET5	196/87	1987	C	Norway	case	+
-	ET5	Mk521/99	1999	B	Ivory Coast	case	+
-	ET5	GR 4/00	2000	-	Greece	case	+
11	ET37	FAM18	1983	C:2a:P1.5,2	USA	case	+
11	ET37	L93/4286	1993	C	UK	case	+
-	ET37	NGP165	1974	B:NT:P1.2	Norway	-	+
-	ET37	M986	1963	B:2a:P1.5,2	USA	case	+
-	ET37	C4678	1998	C:2a:P1.5,2	Germany	case	+
-	ET37	95N477	1995	B:2a:P1.2	Australia	case	-
-	ET37	BRAZ10	1976	C	Brazil	case	+
-	ET37	F1576	1984	C	Ghana	case	+

-	ET37	M597	1988	C	Israel	case	+
-	ET37	500	1984	C	Italy	case	+
-	ET37	D1	1989	C	Mali	case	+
-	ET37	NG P20	1969	B	Norway	case	+
-	ET37	MA-5756	1985	C	Spain	case	+
-	ET37	38VI	1964	B	USA	carrier	+
-	ET37	N1/99	1999	C:2a	Norway	case	+
-	ET37	N28/00	2000	W-135:2a	Norway	case	+
66	A4	973-1720	1997	C:2b:P1.2	Australia	case	+
153	A4	961-5945	1996	B:2b:P1.21,16	Australia	case	+
-	A4	5/99	1999	B:2b:P1.5,2	Norway	case	+
-	A4	312294	1995	C:2b:P1.5,2	UK	case	+
-	A4	96217	1996	B:2b:P1.5,10	Canada	case	+
-	A4	G2136	1986	B	UK	case	+
-	A4	312 901	1996	C	UK	case	+
-	A4	AK22	1992	B	Greece	case	+
-	A4	BZ10	1967	B	Holland	case	+
-	A4	BZ163	1979	B	Holland	case	+
-	A4	B6116/77	1977	B	Iceland	case	+
-	A4	94/155	1994	C	New Zealand	case	+
-	A4	SB25	1990	C	South Africa	case	+
-	A4	N53/00	2000	C:2b:P1.5,2	Norway	case	+
-	A4	N62/00	2000	C:2b:P1.5,2	Norway	case	+
41	Lin.III	BZ198	1986	B:NT	Netherlands	case	-
42	Lin.III	M198/254	1998	B:4:P1.4	New Zealand	case	-
158	Lin.III	972-0319	1997	B:NT:P1.4	Australia	case	-
159	Lin.III	980-2543	1998	B:NT:P1.4	Australia	case	-
1127	Lin.III	67/00	2000	B:4,7	Norway	case	-
-	Lin.III	93/114	1993	C:4:P1.4	Belgium	case	-
-	Lin.III	M198/172	1998	B:4:P1.4	New Zealand	case	-
-	Lin.III	347/97	1997	B:4:P1.4	New Zealand	case	-
-	Lin.III	386/98	1998	B:4:P1.4	New Zealand	case	-
-	Lin.III	389/98	1998	B:4:P1.4	New Zealand	case	-
-	Lin.III	392/98	1998	B:4:P1.4	New Zealand	case	-

-	Lin.III	394/98	1998	B:4:P1.4	New Zealand	case	-
-	Lin.III	400	1991	B	Austria	case	-
-	Lin.III	M40/94	1994	B	Chile	case	-
-	Lin.III	AK50	1992	B	Greece	case	-
-	Lin.III	M-101/93	1993	B	Iceland	case	-
-	Lin.III	931905	1993	B	Netherlands	case	-
-	Lin.III	91/40	1991	B	New Zealand	case	-
-	Lin.III	50/94	1994	B	Norway	case	-
-	Lin.III	N45/96	1996	B	Norway	case	-
-	Lin.III	88/03415	1988	B	Scotland	case	-
1	s I	BZ133	1977	B:NT	Netherlands	case	-
5	s III	F6124	1988	A	Chad	case	+
4	s IV-1	205900	1990	A 4,21:P1.7:1	Mali	case	-
4	s IV-1	Z2491	1983	A	Gambia	case	-
12	other	NG3/88	1988	B:8(2):P1.1	Norway	case	-
13	other	NG6/88	1988	B:NT:P1.1	Norway	case	-
14	other	NGF26	1988	B:NT:P1.16	Norway	carrier	-
15	other	NGE31	1988	B:NT	Norway	carrier	-
18	other	528	1989	B: nd	Russia	case	-
20	other	1000	1988	B: NT:P1.5	Russia	case	-
22	other	A22	1986	W-135	Norway	carrier	-
26	other	NGE28	1988	B:4	Norway	carrier	+
29	other	860800	1986	Y	Netherlands	case	-
31	other	E32	1988	Z	Norway	carrier	-
35	other	SWZ107	1986	B:4:P1.2	Switzerland	case	-
36	other	NGH38	1988	B:NT:P1.3	Norway	carrier	-
38	other	BZ232	1964	B:NT:P1.2	Netherlands	case	-
39	other	E26	1988	X	Norway	carrier	-
43	other	NGH15	1988	B:8:P1.15	Norway	carrier	-
47	other	NGH36	1988	B:8:P1.2	Norway	carrier	-
48	other	BZ147	1963	B:NT	Netherlands	case	-
49	other	297-0	1987	B:4:P1.15	Chile	carrier	-
540	other	2996	1975	B:2b:P1.5-1,2	UK	case	+
1034	other	96/1101	1996	C:14:P.1.1,7	Belgium	case	-

-	other	15	1990	B:14,19:P1.9,15	Slovenia	case	-
-	other	M1090	1996	B:4	Israel	case	-
-	other	M1096	1996	C:NT:P1.5	Israel	case	-
-	other	B3937	1995	B:22:P1.16	Germany	case	+
-	other	31	1993	B:4	Finland	case	-
-	other	95074	1995	B:NT:P1.13	Canada	case	+
-	other	660/94	1994	B:4:P1.6	Algeria	case	-
-	other	30/93	1993	B:14:P1.14	Argentina	case	-
-	other	24370	1996	B:ND	South Africa	case	-
-	other	241175I	1993	NG:21:P1.16	Iceland	carrier	-
-	other	171274I	1993	NG:15:-	Iceland	carrier	-
-	other	65/96	1996	B:4:P1.14	Norway	carrier	+
-	other	66/96	1996	B:17:P1.15	Norway	carrier	-
-	other	149/96	1996	B:1,19:P1.5,2	Belgium	carrier	+
-	other	16060	1991	B:4:P1.14	Belgium	carrier	-
-	other	16489	1991	NG:21:P.1.1	Norway	carrier	-
-	other	16990	1991	NG:14:P1.5,2,6	Norway	carrier	-
-	other	2022	1991	NG:4:P1.10	Norway	carrier	+
-	other	M136	1968	B:11:P1.15	USA	case	-
-	other	860060	1988	X	Holland	case	-
-	other	NG H41	1986	B	Norway	carrier	-
-	other	NG G40	1988	B	Norway	carrier	-
-	other	NG4/88	1988	B	Norway	case	-
-	other	EG 327	1985	B	DDR	case	-
-	other	EG 328	1985	B	DDR	case	-
-	other	3906	1977	B	China	case	-
-	other	NG E30	1988	B	Norway	carrier	-
-	other	71/94	1994	Y	Norway	case	-
-	other	DK24	1940	B	Denmark	case	-
-	-	C11	1965	C :16 :P1.7a,1	Germany	-	+
-	-	pmc8	-	C	-	-	+
-	-	NMB	1968	B:2b:P1.5,2	USA	case	+
-	-	8047	1978	B:2b:P1.2	USA	case	+
-	-	S3446	1972	B:14:P1.23,14	USA	case	-

-	-	ISS 749	1996	B:14:P1.13	Italy	case	-
-	-	ISS 759	1996	C:2b:P1.2	Italy	case	+
-	-	ISS 832	1997	B:15:P1.7	Italy	case	+
-	-	ISS 838	1997	C:2a:P1.5,2	Italy	case	+
-	-	ISS1001	1999	B:14:P1.13	Italy	case	-
-	-	ISS1024	2000	C:2b:P1.5	Italy	case	+
-	-	ISS1026	2000	B:4:P1.13	Italy	case	-
-	-	ISS1071	2000	B:15:P1.7,16	Italy	case	+
-	-	ISS1102	2000	B:15:P1.4	Italy	case	-
-	-	ISS1104	2000	B:15:P1.7,16	Italy	case	+
-	-	ISS1106	2000	B:4:P1.4	Italy	case	-
-	-	ISS1113	2000	C:2a:P1.5	Italy	case	+
-	-	N1002/90	-	-	Brazil	-	+
-	-	IMC2135	-	-	Brazil	-	+
-	-	NM001	-	B:4:P1.4	UK	case	-
-	-	NM002	-	B:NT:P1.16	UK	case	-
-	-	NM004	-	B:NT:P1.14	UK	case	-
-	-	NM008	-	B:4:P1.4	UK	case	-
-	-	NM009/10	-	B:4:P1.3,6	UK	case	-
-	-	NM021	-	B:4:P1.16	UK	case	-
-	-	NM036	-	C:2a:P1.10	UK	case	+
-	-	NM037	-	B:2b:P1.10	UK	case	+
-	-	NM050	-	B:NT:P1.9	UK	case	-
-	-	NM058	-	B:NT:NST	UK	case	-
-	-	NM066	-	B:15:P1.7,16	UK	case	+
-	-	NM067	-	C:2a:NST	UK	case	+
-	-	NM069	-	B:15:P1.7,16	UK	case	+
-	-	NM081	-	C:2a:P1.5,2	UK	case	+
-	-	NM088	-	C:2a:P1.5,2	UK	case	+
-	-	NM092	-	B:4:P1.4	UK	case	-
-	-	NM106	-	B:NT:P1.4	UK	case	-
-	-	NM107/8	-	B:4:P1.4	UK	case	-
-	-	NM117	-	B:21:P1.9	UK	case	-
-	-	NM119	-	B:4:P1.15	UK	case	+

-	-	NM131	-	B	UK	case	-
-	-	NM145	-	C	UK	case	+
-	-	NM154	-	C:NT:P1.5,2	UK	case	+
-	-	NM156	-	B:15:P1.16	UK	case	+
-	-	NM167	-	B	UK	case	-
-	-	NM184	-	B:NT:P1.5,2	UK	case	-
-	-	NM186	-	B	UK	case	-
-	-	NM188	-	B	UK	case	+
-	-	NM200	-	B:4:P1.4	UK	case	-

TABLE III — SEQUENCE LISTING

SEQ ID NO:	Description
1	allele 1 of 961
2	allele 2 of 961
3	allele 3 of 961
4	allele 1 of 961 (first-ATG start)
5	allele 2 of 961 (first-ATG start)
6	allele 3 of 961 (first-ATG start)
7	variant allele 2 of 961 in strain ISS1024
8	variant allele 2 of 961 (first-ATG start) in strain ISS1024
9	variant allele 3 of 961 in strains 973-1720 and ISS759
10	variant allele 3 of 961 (first-ATG start) in strains 973-1720 and ISS759
11	961 allele 1/2 chimera (strain 95330)
12	961 allele 1/2 chimera (strain 95330) (first-ATG start)
13	961 allele C
14	961 allele C (first-ATG start)
15	coding sequence for SEQ ID 13
16-31	PCR primers
32	SEQ ID 650 from WO99/24578
33-39	Domain derivatives of SEQ ID 32

CLAIMS

1. A protein comprising the amino acid sequence of one or more of SEQ IDs 1 to 14.

2. A protein comprising an amino acid sequence having at least 50% sequence identity to one or more of SEQ IDs 1 to 14.

3. A protein comprising a fragment of one or more of SEQ IDs 1 to 14.

4. The protein of claim 1, claim 2 or claim 3, including the heptad sequence $(AA_1AA_2AA_3AA_4AA_5AA_6AA_7)_r$ wherein: AA_1 is Leu, Ile, Val or Met; each of AA_2 AA_3 AA_4 AA_5 AA_6 and AA_7 may independently be any amino acid; and *r* is an integer of 1 or more

5. The protein of claim 3, comprising amino acids 24 to 351 of SEQ ID 3.

6. Nucleic acid encoding a protein any one of claims 1 to 5.

7. An immunogenic composition comprising (a) a Neisserial NadA protein and/or (b) nucleic acid encoding a Neisserial NadA protein.

8. A method for raising an antibody response in a mammal, comprising administering the immunogenic composition of claim 7 to the mammal.

9. A method for protecting a mammal against a Neisserial infection, comprising administering to the mammal the immunogenic composition of claim 7.

10. The use of a Neisserial NadA protein, or nucleic acid encoding NadA protein, in the manufacture of a medicament for preventing Neisserial infection in a mammal.

11. Neisserial NadA protein, or nucleic acid encoding NadA protein, for use as a medicament.

12. The method, use or protein of any one of claims 8 to 111, wherein the NadA protein comprises the amino acid sequence of one or more of SEQ IDs 1 to 12.

13. The method, use or protein of any one of claims 8 to 12, wherein the Neisserial infection is an infection with *N.meningitidis* from hypervirulent lineages ET-5, EY-37 and cluster A4.

14. The composition of claim 7, additionally comprising an antigen which, when administered to a mammal, elicits an immune response which is protective against a lineage III strain of *N.meningitidis*.

15. A method for purifying processed App protein, comprising the steps of: expressing a gene encoding App protein in a non-Neisserial host cell; and purifying processed App protein from the culture medium.

16. The method of claim 15, wherein the non-Neisserial host cell is *E.coli*.

17. Purified protein obtainable by the process of claim 15 or claim 16.

18. A method for preventing the attachment of a Neisserial cell to an epithelial cell, wherein the ability of one or more App, ORF40 and/or NadA to bind to the epithelial cell is blocked.

19. The method of claim 18, wherein the binding is blocked using an antibody specific for App, ORF40 and/or NadA.

20. A method for preventing the attachment of a Neisserial cell to an epithelial cell, wherein protein expression from one or more of App, ORF40 and/or NadA is inhibited.

21. The method of claim 20, wherein expression is inhibited using antisense techniques.

22. Nucleic acid comprising a fragment of *x* or more nucleotides from nucleic acid which encodes App, ORF40 or NadA, wherein *x* is at least 8.

23. The nucleic acid of claim 22 having formula 5'–$(N)_a$–(X)–$(N)_b$–3', wherein 0>a>15, 0>b>15, N is any nucleotide, and X is a fragment of at least 8 contiguous nucleotides from a nucleic acid which encodes App, ORF40 or NadA.

24. A *Neisseria* bacterium in which one or more of App, ORF40 and/or NadA has been knocked out.

25. A method for preventing the attachment of a Neisserial cell to an epithelial cell, wherein one or more of App, ORF40 and/or NadA has a mutation which inhibits its activity.

26. A mutant protein, wherein the mutant protein comprises the amino acid sequence of App, ORF40 and/or NadA, or a fragment thereof, but wherein one or more amino acids of said amino acid sequence is/are mutated.

27. Nucleic acid encoding the mutant protein of claim 26.

28. A method for producing the nucleic acid of claim 27, comprising the steps of: (a) providing source nucleic acid encoding App, ORF40 or NadA, and (b) performing mutagenesis on said source nucleic acid to provide nucleic acid encoding a mutant protein.

29. A method for screening for compounds which inhibit the binding of a Neisserial cell to an epithelial cell, comprising the steps of (a) incubating App, ORF40 and/or NadA protein with an epithelial cell and a test compound; (b) testing the mixture to determine if the interaction between the protein and the epithelial cell has been inhibited.

30. A compound identified using the method of claim 29.

31. A composition comprising (a) an *E.coli* bacterium which expresses App and/or ORF40 (and, optionally, NadA) and (b) an epithelial cell.

32. A method for preparing an outer membrane vesicle (OMV) from a non-Neisserial host cell, characterised in that said cell expresses a gene encoding App, ORF40 or NadA protein.

33. A method for preparing an OMV from a non-Neisserial host cell, characterised in that said cell expresses a gene encoding one or more of the following proteins:

(A) Even SEQ IDs 2-892 from WO99/24578;

(B) Even SEQ IDs 2-90 from WO99/36544;

(C) Even SEQ IDs 2-3020 from WO99/57280;

(D) Even SEQ IDs 3040-3114 from WO99/57280;

(E) SEQ IDs 3115-3241 from WO99/57280;

(F) The 2160 proteins NMB0001 to NMB2160 from Tettelin et al. [supra];

(G) A protein comprising the amino acid sequence of one or more of (A) to (F);

(H) A protein sharing sequence identity with the amino acid sequence of one or more of (A) to (F); and

(I) A protein comprising a fragment of one or more of (A) to (F).

34. The method of claim 32 or claim 33, wherein the non-Neisserial host cell is *E.coli*.

35. OMVs obtainable by the process of claim 32, claim 33 or claim 34.

36. An outer membrane vesicle from a non-Neisserial host cell, characterised in that said cell expresses a gene encoding App, ORF40 or NadA protein.

37. An outer membrane vesicle from a non-Neisserial host cell, characterised in that said cell expresses a gene encoding one or more of proteins (A) to (I) as defined in claim 33.

38. A protein comprising the amino acid sequence of App, except that amino acid Asp-158, His-115 and/or Ser-267 is mutated.

39. The protein of claim 38, wherein Ser-267, Asp-158 or His-115 is replaced with one of the 19 other naturally-occurring amino acids.

40. A protein comprising the amino acid sequence of App, except that one or more amino acid(s) between Ser-1064 and Arg-1171 is mutated.

41. The protein of claim 40, wherein the mutation is a deletion, an insertion, a truncation or a substitution.

42. The protein of claim 40 or claim 41, wherein the residue which is mutated is S-1064, D-1065, K-1066, L-1067, G-1068, K-1069, A-1070, E-1071, A-1072, K-1073, K-1074, Q-1075, A-1076, E-1077, K-1078, D-1079, N-1080, A-1081, Q-1082, S-1083, L-1084, D-1085, A-1086, L-1087, I-1088, A-1089, A-1090, G-1091, R-1092, D-1093, A-1094, V-1095, E-1096, K-1097, T-1098, E-1099, S-1100, V-1101, A-1102, E-1103, P-1104, A-1105, R-1106, Q-1107, A-1108, G-1109, G-1110, E-1111, N-1112, V-1113, G-1114, I-1115, M-1116, Q-1117, A-1118, E-1119, E-1120, E-1121, K-1122, K-1123, R-1124, V-1125, Q-1126, A-1127, D-1128, K-1129, D-1130, T-1131, A-1132, L-1133, A-1134, K-1135, Q-1136, R-1137, E-1138, A-1139, E-1140, T-1141, R-1142, P-1143, A-1144, T-1145, T-1146, A-1147, F-1148, P-1149, R-1150, A-1151, R-1152, R-1153, A-1154, R-1155, R-1156, D-1157, L-1158, P-1159, Q-1160, L-1161, Q-1162, P-1163, Q-1164, P-1165, Q-1166, P-1167, Q-1168, P-1169, Q-1170 and/or R-1171.

43. A protein comprising the amino acid sequence of App, except that one or more of amino acids Phe-956, Asn-957, Ala-1178 & Asn-1179 (numbered according to SEQ ID 32 herein) is mutated.

44. The protein of claim 43, wherein the mutation is a deletion, an insertion, a truncation or a substitution.

45. A protein comprising the amino acid sequence of a processed App, wherein said processed App does not comprise the C-terminus domain which is downstream of an autoproteloytic cleavage site in full-length App.

46. The protein of claim 45, comprising one or more of SEQ IDs 33 to 36.

47. A protein comprising the amino acid sequence of a processed App, wherein the C-terminus of said processed App is Phe-956 (numbered according to SEQ ID 32 herein).

48. A protein comprising the amino acid sequence of a processed App, wherein the C-terminus of said processed App is Ala-1178 (numbered according to SEQ ID 32 herein).

49. A protein comprising an amino acid sequence selected from the group consisting of SEQ IDs 33, 34, 35, 36, 37, 38 & 39.

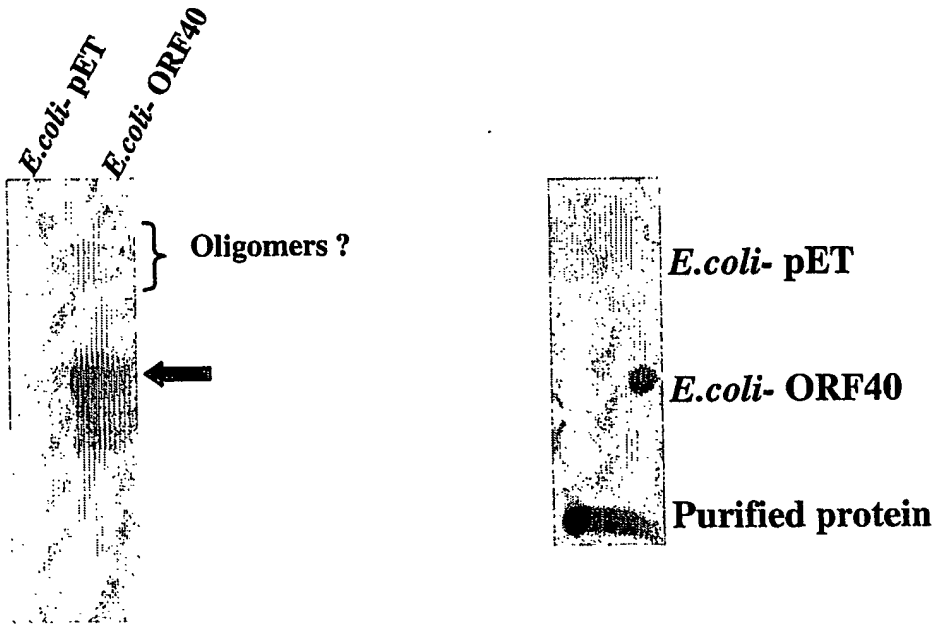
50. A protein comprising an amino acid sequence with at least 50% sequence identity to one or more of SEQ IDs 33, 34, 35, 36, 37, 38 & 39.

51. A protein comprising a fragment of one or more of SEQ IDs 33, 34, 35, 36, 37, 38 & 39.

52. Nucleic acid encoding the protein of any one of claims 38 to 51.

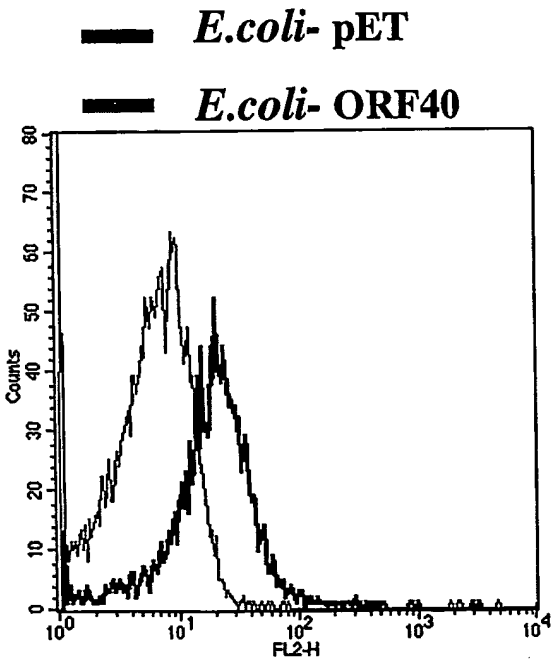
53. A method for producing the nucleic acid of claim 52, comprising the steps of: (a) providing source nucleic acid encoding App, ORF40 or NadA, and (b) performing mutagenesis on said source nucleic acid to provide nucleic acid encoding a protein of any one of claims 38 to 51.

FIGURE 1



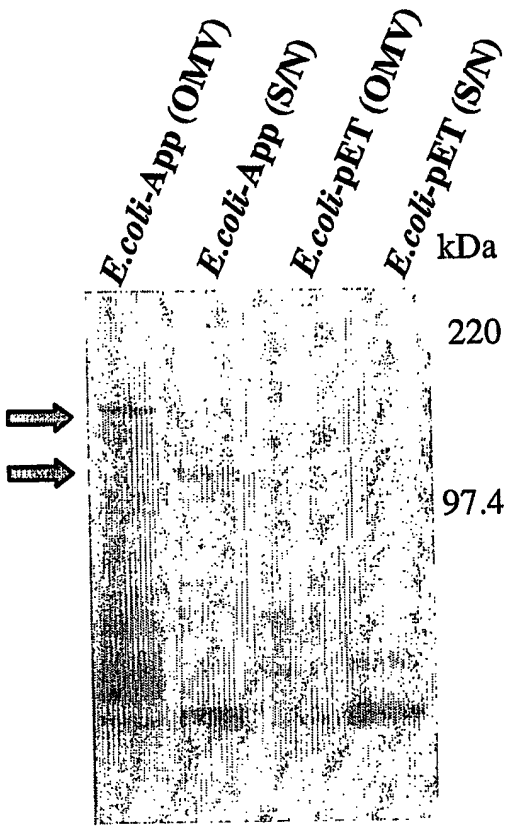
Whole cells
Western blot

Whole-cell
immunoblot

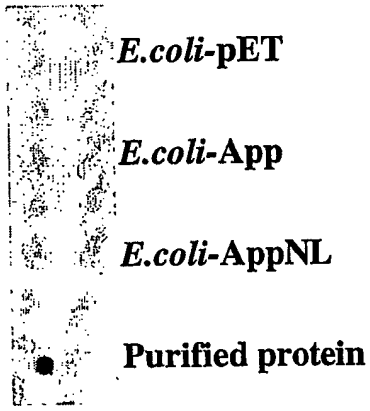


FACS

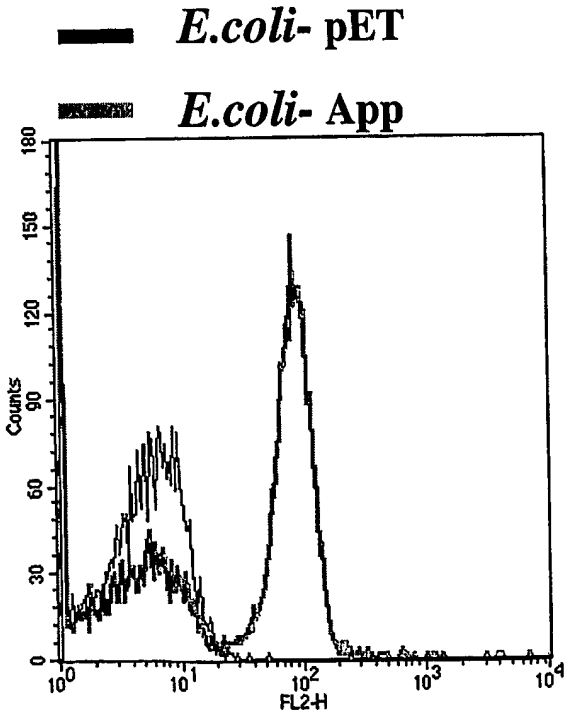
FIGURE 2



Western blot

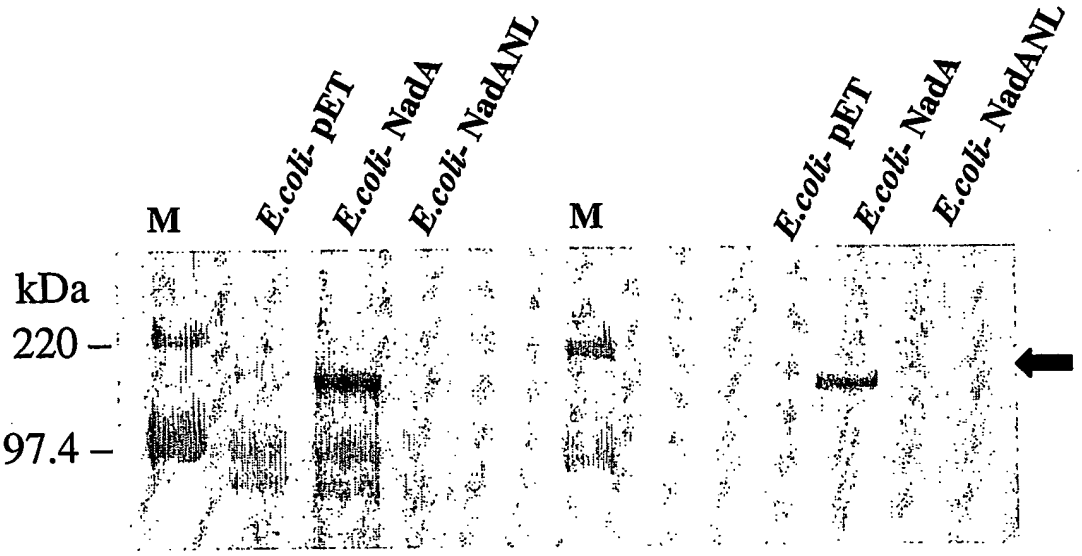


Whole-cell immunoblot

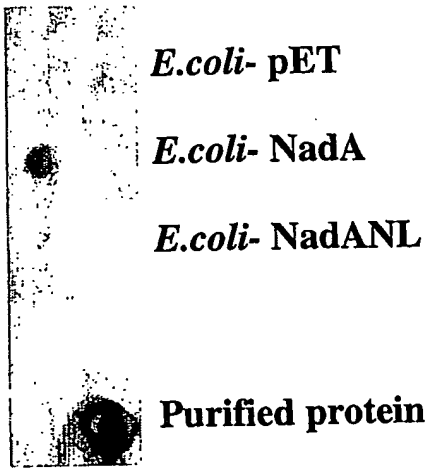


FACS

FIGURE 3



SDS-PAGE



Whole-cell immunoblot

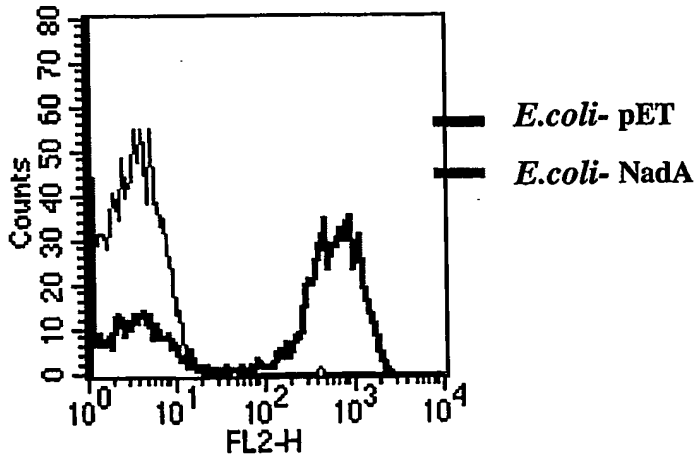


FIGURE 4

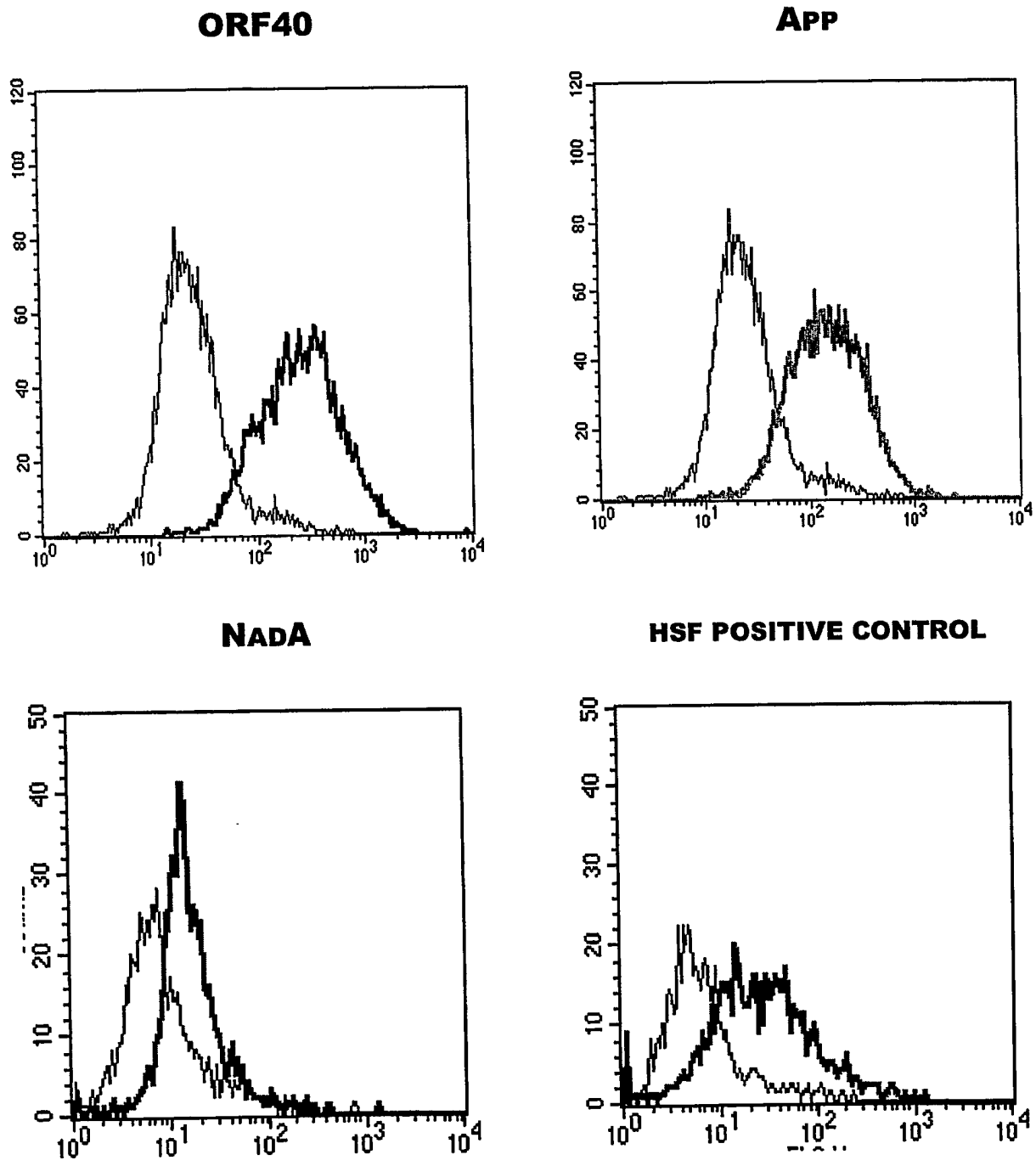


FIGURE 5

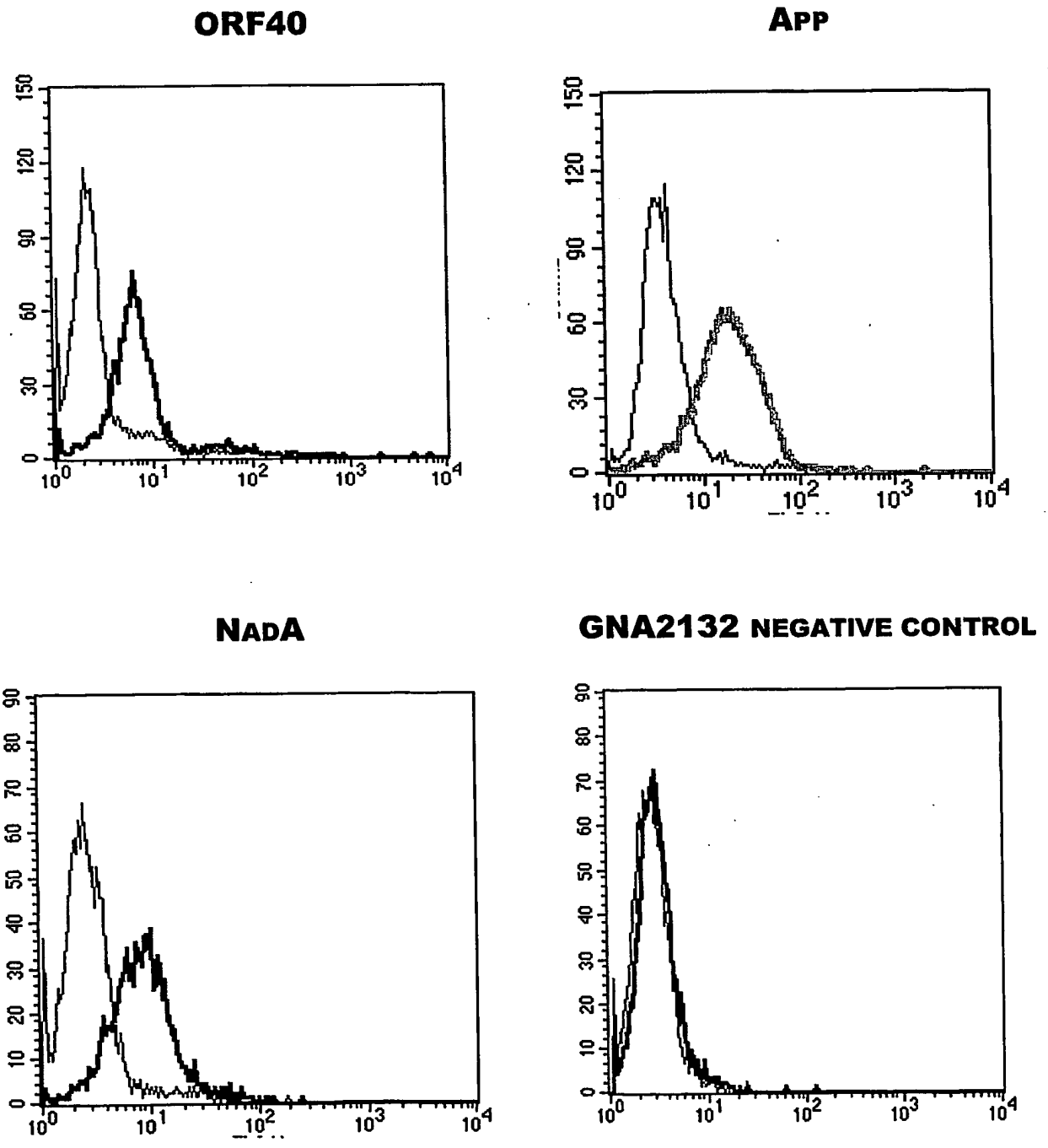


FIGURE 6A

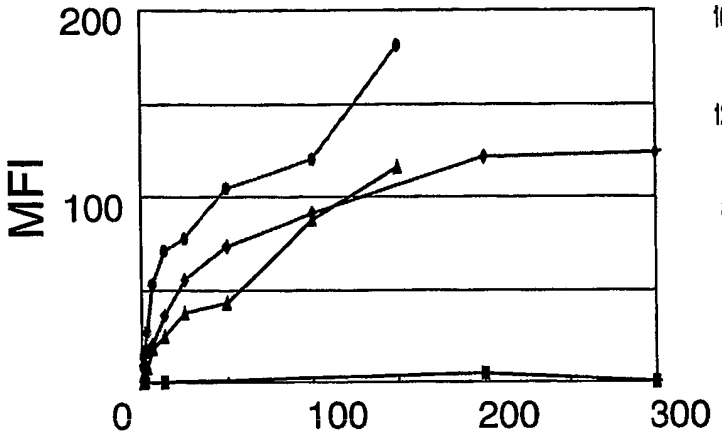


FIGURE 6B

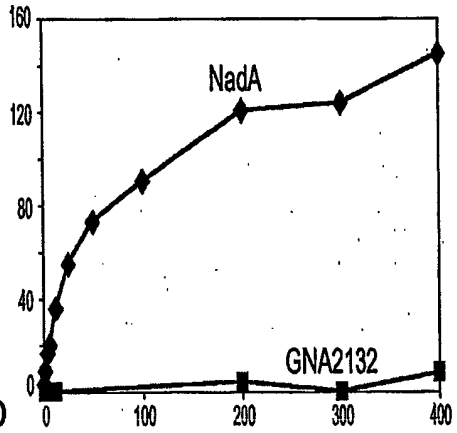


FIGURE 7

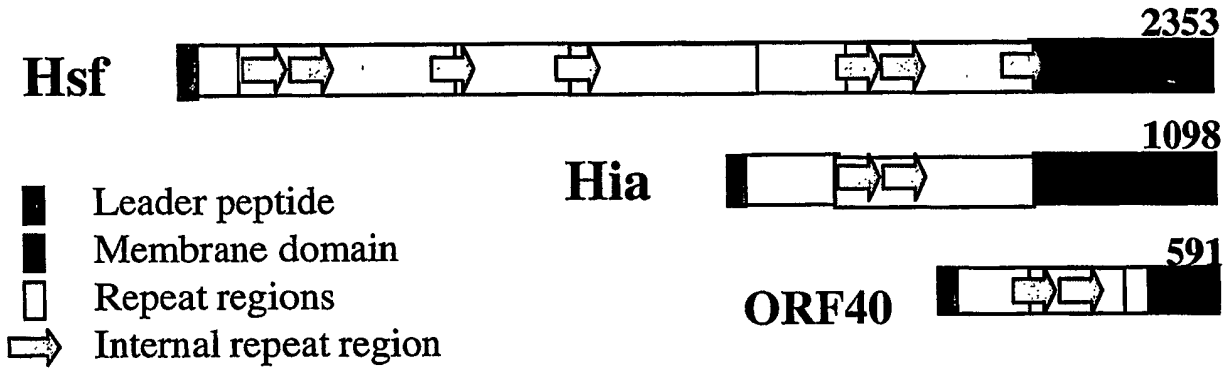


FIGURE 8

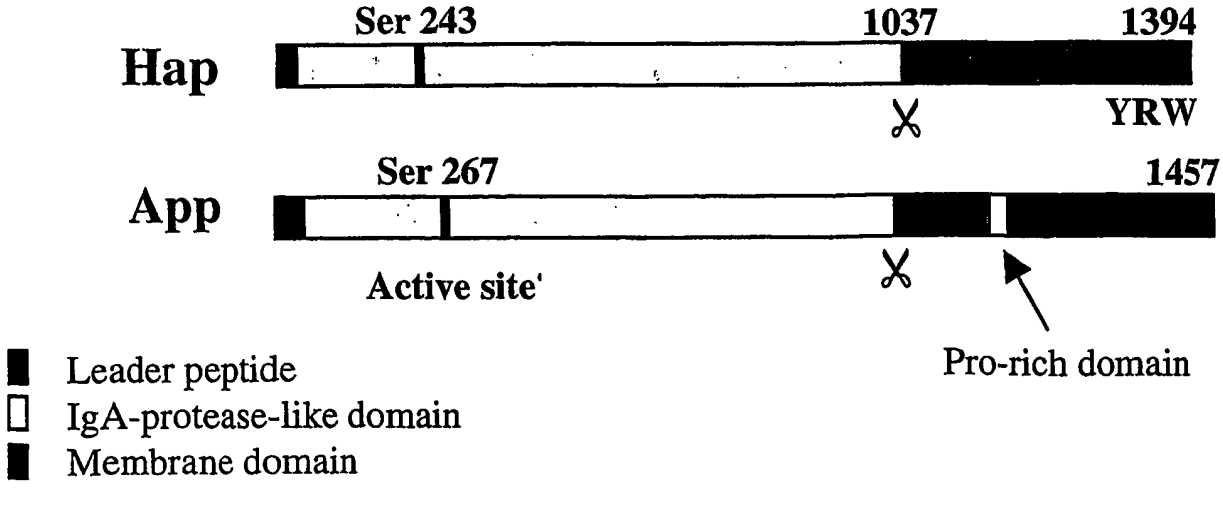


FIGURE 9A

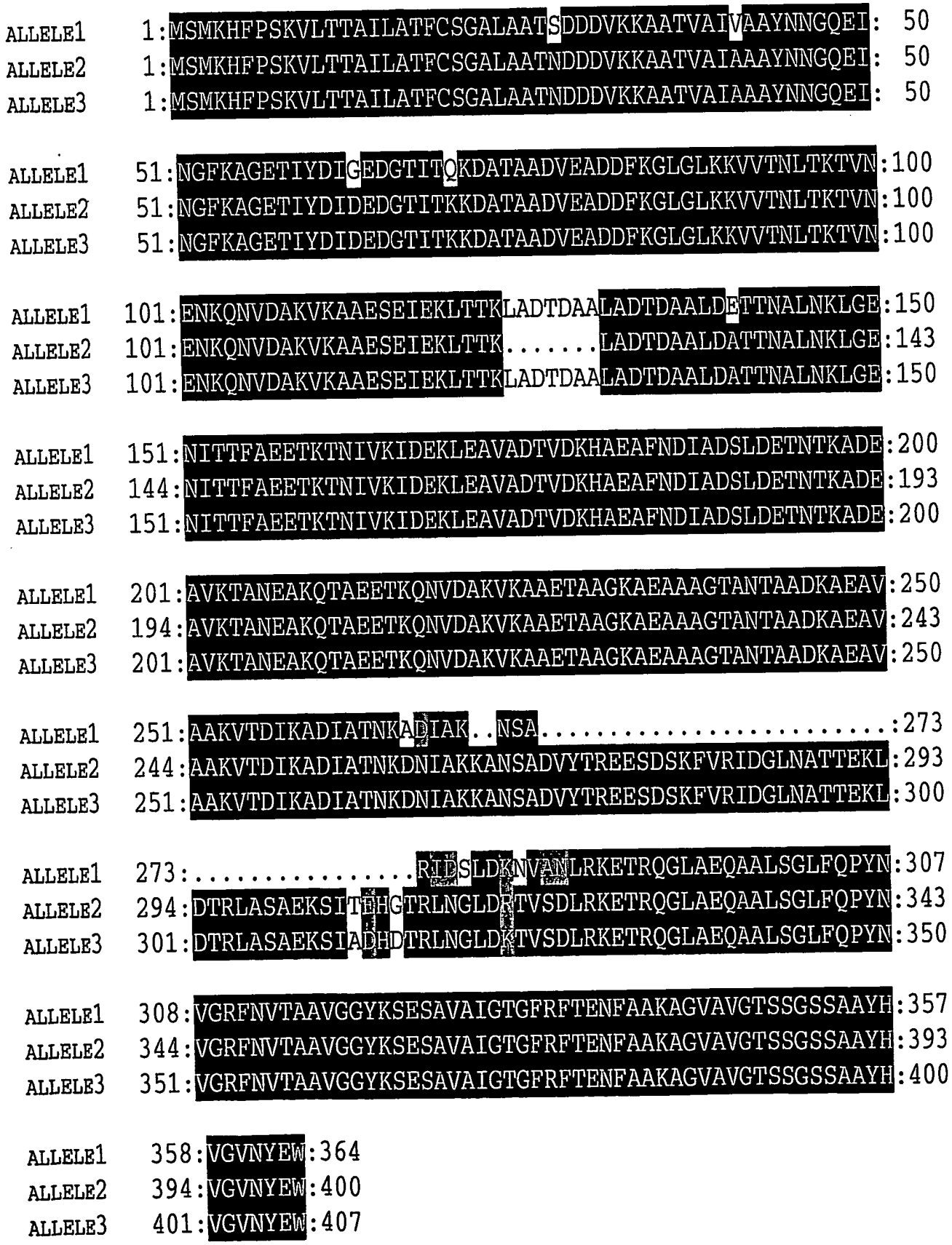
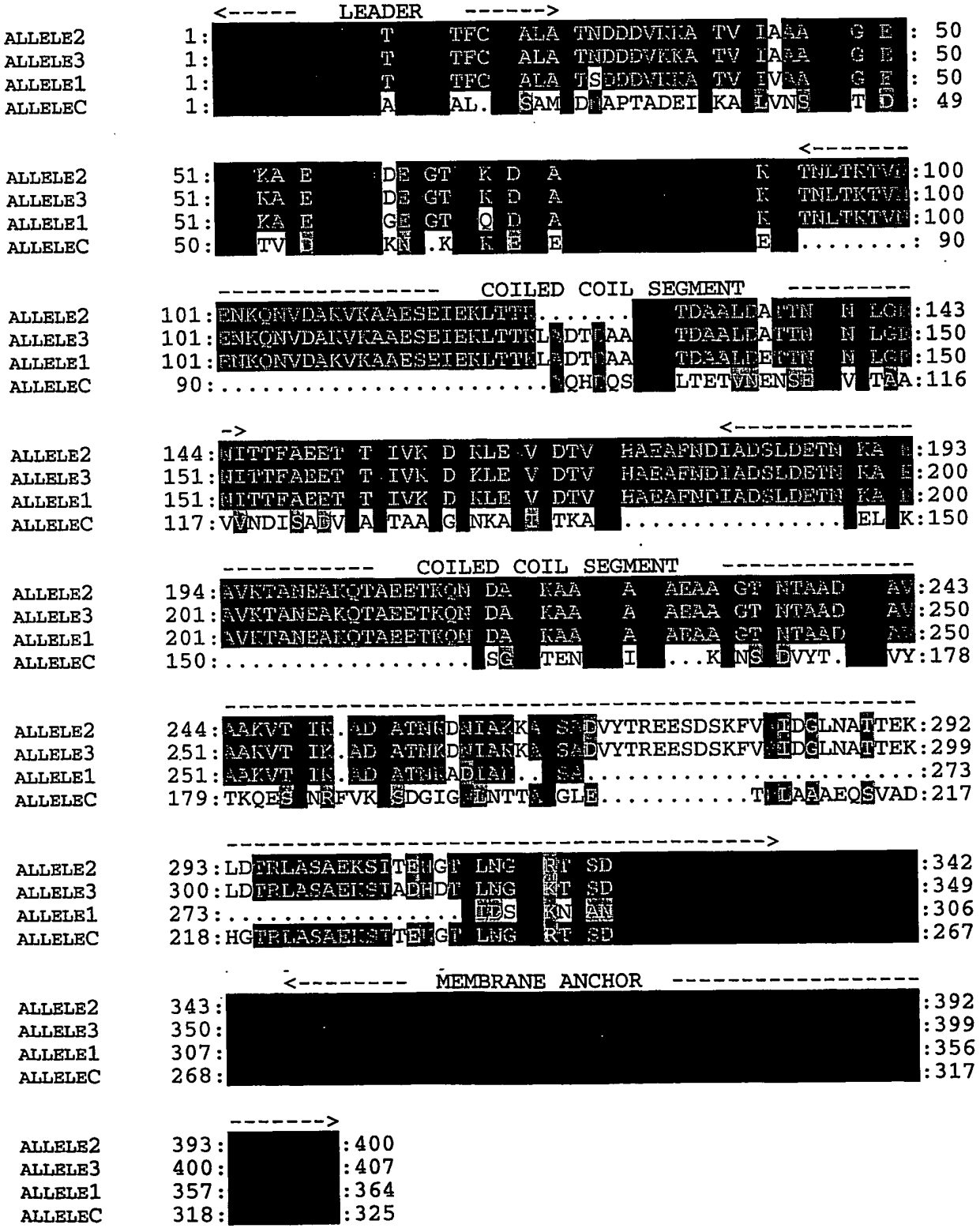

ALLELE1 1:MSMKHFPSKVLTTAILATFCSGALAATSDDDVKKAATVAIVAAYNNGQEI: 50
ALLELE2 1:MSMKHFPSKVLTTAILATFCSGALAATNDDDVKKAATVAIAAAYNNGQEI: 50
ALLELE3 1:MSMKHFPSKVLTTAILATFCSGALAATNDDDVKKAATVAIAAAYNNGQEI: 50
ALLELE1 51:NGFKAGETIYDIGEDGTITQKDATAADVEADDFKGLGLKKVVTNLTKTVN:100
ALLELE2 51:NGFKAGETIYDIDEDGTITKKDATAADVEADDFKGLGLKKVVTNLTKTVN:100
ALLELE3 51:NGFKAGETIYDIDEDGTITKKDATAADVEADDFKGLGLKKVVTNLTKTVN:100
ALLELE1 101:ENKQNVDAKVKAAESEIEKLTTKLADTDAALADTDAALDETTNALNKLGE:150
ALLELE2 101:ENKQNVDAKVKAAESEIEKLTTK.....LADTDAALDATTNALNKLGE:143
ALLELE3 101:ENKQNVDAKVKAAESEIEKLTTKLADTDAALADTDAALDATTNALNKLGE:150
ALLELE1 151:NITTFAEETKTNIVKIDEKLEAVADTVDKHAEAFNDIADSLDETNTKADE:200
ALLELE2 144:NITTFAEETKTNIVKIDEKLEAVADTVDKHAEAFNDIADSLDETNTKADE:193
ALLELE3 151:NITTFAEETKTNIVKIDEKLEAVADTVDKHAEAFNDIADSLDETNTKADE:200
ALLELE1 201:AVKTANEAKQTAEETKQNVDAKVKAAETAAGKAEAAAGTANTAADKAEAV:250
ALLELE2 194:AVKTANEAKQTAEETKQNVDAKVKAAETAAGKAEAAAGTANTAADKAEAV:243
ALLELE3 201:AVKTANEAKQTAEETKQNVDAKVKAAETAAGKAEAAAGTANTAADKAEAV:250
ALLELE1 251:AAKVTDIKADIATNKADIAK..NSA.....:273
ALLELE2 244:AAKVTDIKADIATNKDNIAKKANSADVYTREESDSKFVRIDGLNATTEKL:293
ALLELE3 251:AAKVTDIKADIATNKDNIAKKANSADVYTREESDSKFVRIDGLNATTEKL:300
ALLELE1 273:.....RIDSLDKNVANLRKETRQGLAEQAALSGLFQPYN:307
ALLELE2 294:DTRLASAEKSITEHGTRLNGLDRTVSDLRKETRQGLAEQAALSGLFQPYN:343
ALLELE3 301:DTRLASAEKSIADHDTRLNGLDKTVSDLRKETRQGLAEQAALSGLFQPYN:350
ALLELE1 308:VGRFNVTAAVGGYKSESAVAIGTGFRFTENFAAKAGVAVGTSSGSSAAYH:357
ALLELE2 344:VGRFNVTAAVGGYKSESAVAIGTGFRFTENFAAKAGVAVGTSSGSSAAYH:393
ALLELE3 351:VGRFNVTAAVGGYKSESAVAIGTGFRFTENFAAKAGVAVGTSSGSSAAYH:400
ALLELE1 358:VGVNYEW:364
ALLELE2 394:VGVNYEW:400
ALLELE3 401:VGVNYEW:407

FIGURE 9B

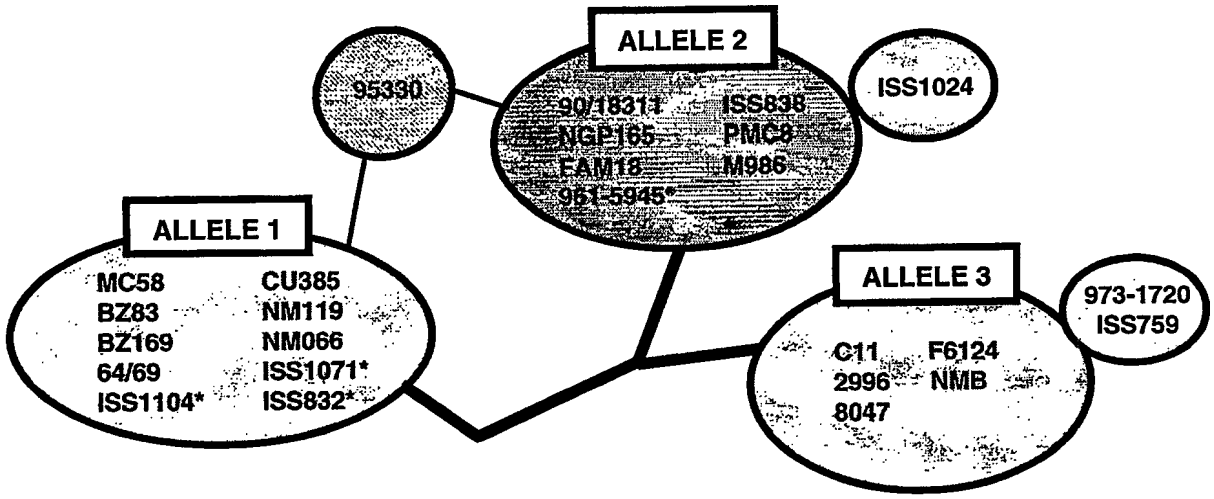

<----- Leader ----->
ALLELE_3 MKHFPSKVLTTAILATFCSGALAATNDDDVKKAATVAIAAAYNNGQEING 50
ALLELE_2 .....
ALLELE_1 .....S.....V.....
<---------
ALLELE 3 FKAGETIYDIDEDGTITKKDATAADVEADDFKGLGLKKVVTNLTKTVNEN 100
ALLELE 2 .....
ALLELE 1 .....G.....Q.....
------- coiled coil segment ---------->
ALLELE_3 KQNVDAKVKAAESEIEKLTTKLADTDAALADTDAALDATTNALNKLGENI 150
ALLELE_2 .....
ALLELE_1 .....E.....
<----------
ALLELE 3 TTFAEETKTNIVKIDEKLEAVADTVDKHAEAFNDIADSLDETNTKADEAV 200
ALLELE 2 .....
ALLELE 1 .....
---------- coiled coil segment ----------
ALLELE_3 KTANEAKQTAEETKQNVDAKVKAAETAAGKAEAAAGTANTAADKAEAVAA 250
ALLELE_2 .....
ALLELE_1 .....
ALLELE 3 KVTDIKADIATNKDNIAKKANSADVYTREESDSKFVRIDGLNATTEKLDT 300
ALLELE 2 .....
ALLELE 1 .....AD...
---------->
ALLELE_3 RLASAEKSIADHDTRLNGLDKTVSDLRKETRQGLAEQAALSGLFQPYNVG 350
ALLELE_2 .....TE.G.....R.....
ALLELE_1 .....IDS...N.AN.....
<---------- membrane anchor ----------
ALLELE 3 RFNVTAAVGGYKSESAVAIGTGFRFTENFAAKAGVAVGTSSGSSAAYHVG 400
ALLELE 2 .....
ALLELE 1 .....
---->
ALLELE_3 VNYEW 405
ALLELE_2 .....
ALLELE_1 .....

FIGURE 9C


<------ LEADER ------>
ALLELE2 1: : 50
ALLELE3 1: : 50
ALLELE1 1: : 50
ALLELEC 1: : 49
<-------
ALLELE2 51: :100
ALLELE3 51: :100
ALLELE1 51: :100
ALLELEC 50: : 90
---------- COILED COIL SEGMENT ---------
ALLELE2 101: :143
ALLELE3 101: :150
ALLELE1 101: :150
ALLELEC 90: :116
-> <----------
ALLELE2 144: :193
ALLELE3 151: :200
ALLELE1 151: :200
ALLELEC 117: :150
---------- COILED COIL SEGMENT ----------
ALLELE2 194: :243
ALLELE3 201: :250
ALLELE1 201: :250
ALLELEC 150: :178

ALLELE2 244: :292
ALLELE3 251: :299
ALLELE1 251: :273
ALLELEC 179: :217
---------->
ALLELE2 293: :342
ALLELE3 300: :349
ALLELE1 273: :306
ALLELEC 218: :267
<-------- MEMBRANE ANCHOR ----------
ALLELE2 343: :392
ALLELE3 350: :399
ALLELE1 307: :356
ALLELEC 268: :317
------->
ALLELE2 393: :400
ALLELE3 400: :407
ALLELE1 357: :364
ALLELEC 318: :325

FIGURE 10




```
  1 MSMKHFPSKVLTTAILATFCSGALAATSDDDVKKAATVAIVAAYNNGQEINGFKAGETIYDIGEDGTITQKDATAADVEADDFKGLGLKKVVTNLTKTVN
  1 MSMKHFPSKVLTTAILATFCSGALAATNDDDVKKAATVAIAAAYNNGQEINGFKAGETIYDIDEDGTITKKDATAADVEADDFKGLGLKKVVTNLTKTVN
  1 MSMKHFPSKVLTTAILATFCSGALAATNDDDVKKAATVAIAAAYNNGQEINGFKAGETIYDIDEDGTITKKDATAADVEADDFKGLGLKKVVTNLTKTVN

101 ENKQNVDAKVKAAESEIEKLTTKLADTDAALADTDAALDETTNALNKLGENITTFAEETKTNIVKIDEKLEAVADTVDKHAEAFNDIADSLDETNTKADE
101 ENKQNVDAKVKAAESEIEKLTTK-------LADTDAALDATTNALNKLGENITTFAEETKTNIVKIDEKLEAVADTVDKHAEAFNDIADSLDETNTKADE
101 ENKQNVDAKVKAAESEIEKLTTKLADTDAALADTDAALDATTNALNKLGENITTFAEETKTNIVKIDEKLEAVADTVDKHAEAFNDIADSLDETNTKADE

201 AVKTANEAKQTAEETKQNVDAKVKAAETAAGKAEAAAGTANTAADKAEAVAAKVTDIKADIATNKADIAK--NSA------------------------
194 AVKTANEAKQTAEETKQNVDAKVKAAETAAGKAEAAAGTANTAADKAEAVAAKVTDIKADIATNKDNIAKKANSADVYTREESDSKFVRIDGLNATTEKL
201 AVKTANEAKQTAEETKQNVDAKVKAAETAAGKAEAAAGTANTAADKAEAVAAKVTDIKADIATNKDNIAKKANSADVYTREESDSKFVRIDGLNATTEKL

274 ----------------RIDSLDKNVANLRKETRQGLAEQAALSGLFQPYNVGRFNVTAAVGGYKSESAVAIGTGFRFTENFAAKAGVAVGTSSGSSAAYH
294 DTRLASAEKSITEHGTRLNGLDRTVSDLRKETRQGLAEQAALSGLFQPYNVGRFNVTAAVGGYKSESAVAIGTGFRFTENFAAKAGVAVGTSSGSSAAYH
301 DTRLASAEKSIADHDTRLNGLDKTVSDLRKETRQGLAEQAALSGLFQPYNVGRFNVTAAVGGYKSESAVAIGTGFRFTENFAAKAGVAVGTSSGSSAAYH

358 VGVNYEW
394 VGVNYEW
401 VGVNYEW
```

FIGURE 11

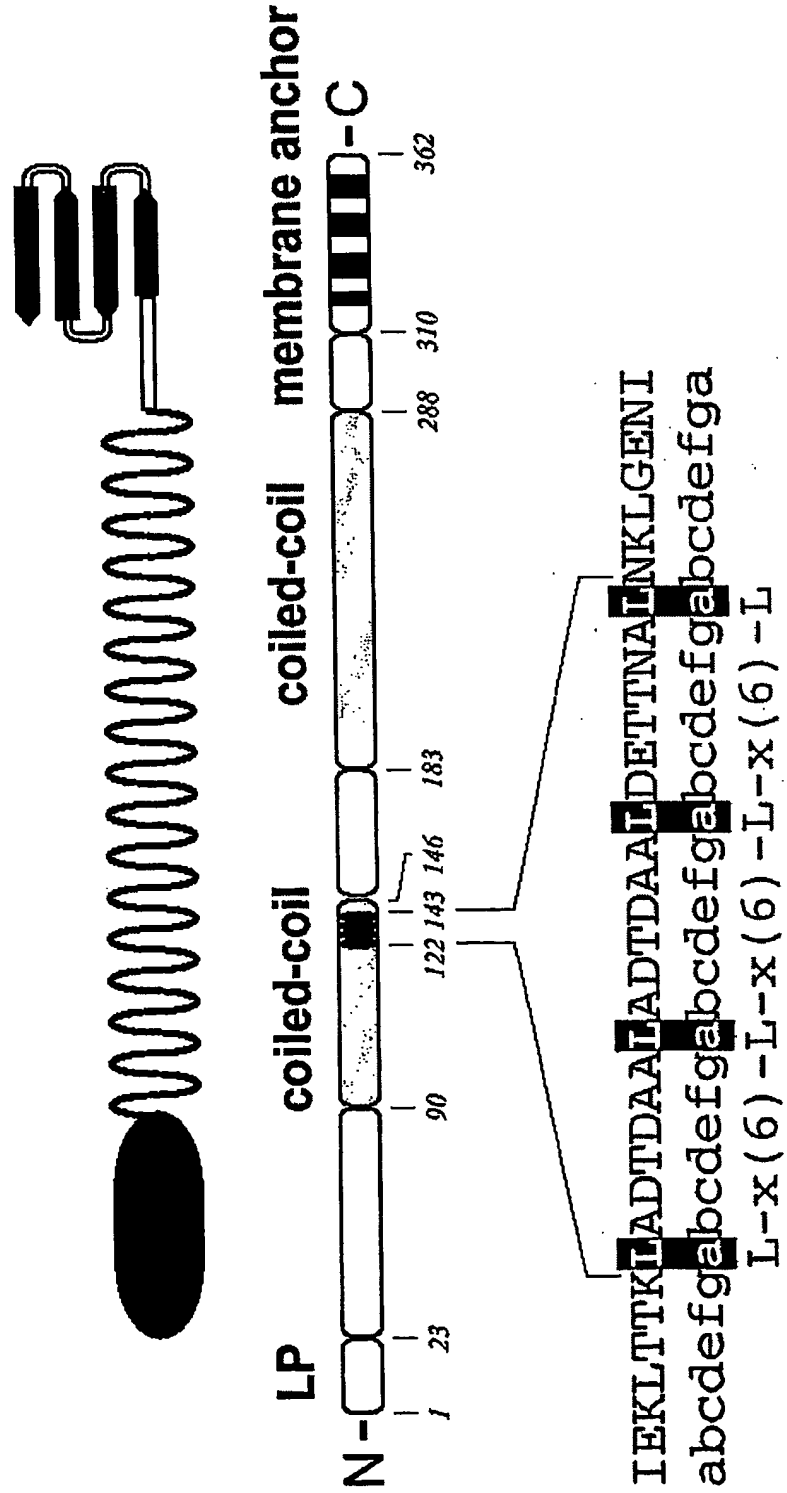

N-
LP
coiled-coil
coiled-coil
membrane anchor
-C
1
23
90
122 143 146
183
288
310
362
IEKLTTKLADTDAALADTDAALDETTNALNKLGENI
abcdefgabcdefgabcdefgabcdefgabcdefga
L-x(6)-L-x(6)-L-x(6)-L

FIGURE 12

ACCCATATCCTGACAAAATTAAGACACGACACCGGCAGAATTGACATCAGCATAATATGC

-35

ACATATTAACAGATATTAATGCCGAACTACCTAACTGCAAGAATTAAATAAATAAATAAA

-10

TAAATAAATAAATAAATAAATTGCGACAATGTATTGTATATATGCCTCCTTTCATATATA

CTTTAATATGTAAACAAACTTGGTGGGGATAAAATACTTACAAAAGATTTCCGCCCCATT

-35 -10

TTTTATCCACTCACAAAGGTAATGAGCATG

RBS

FIGURE 12A

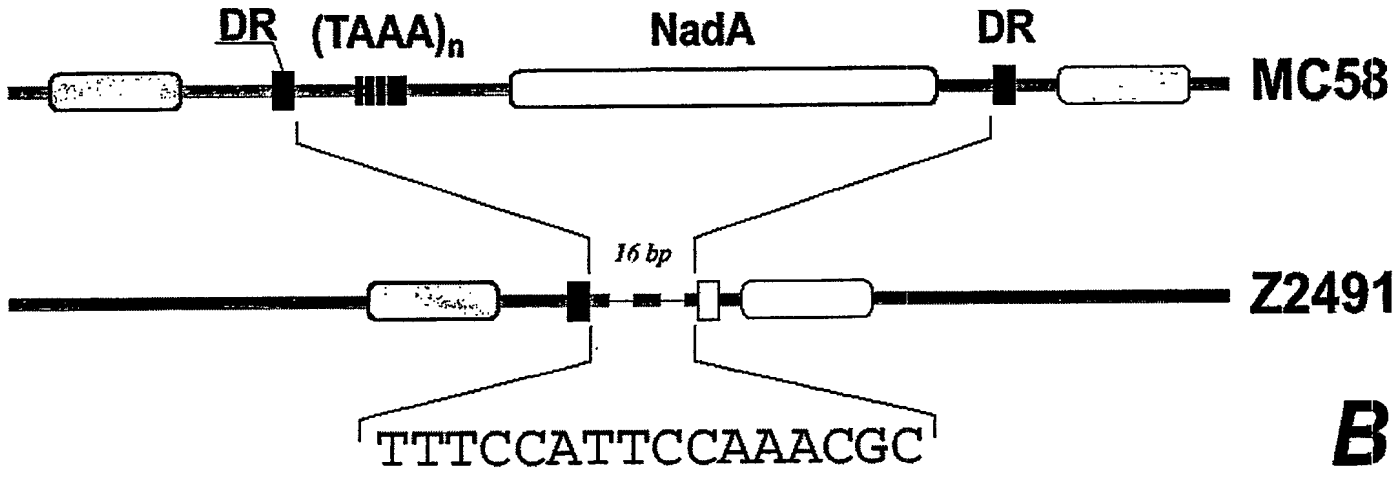


FIGURE 12B

FIGURE 13

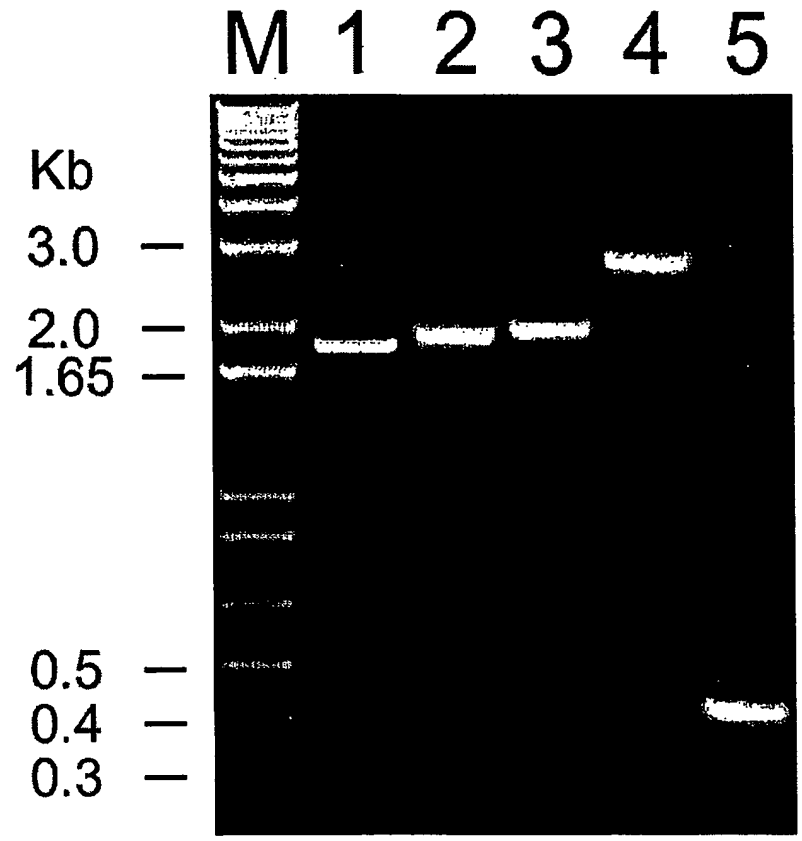


FIGURE 14

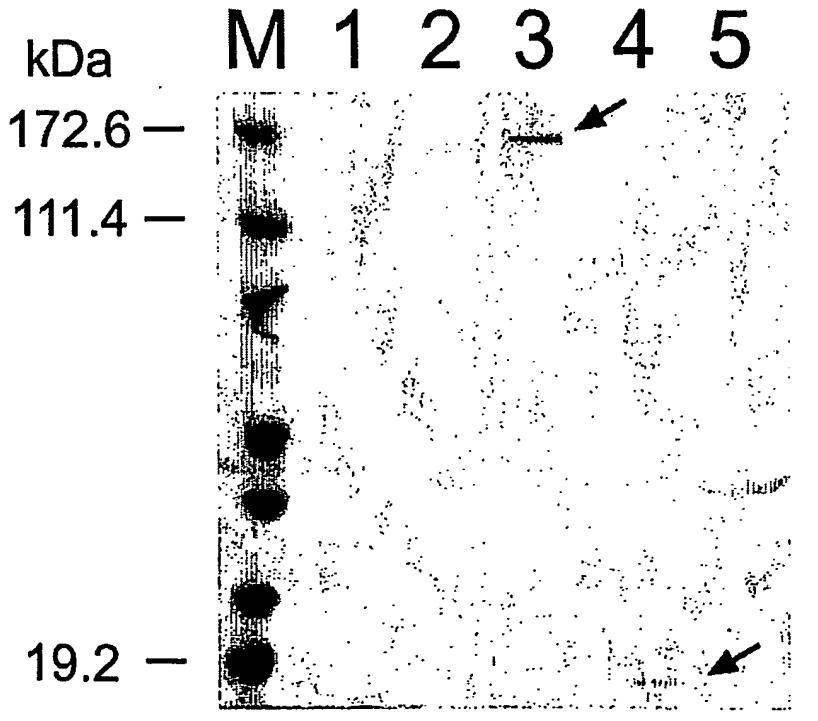


FIGURE 15A

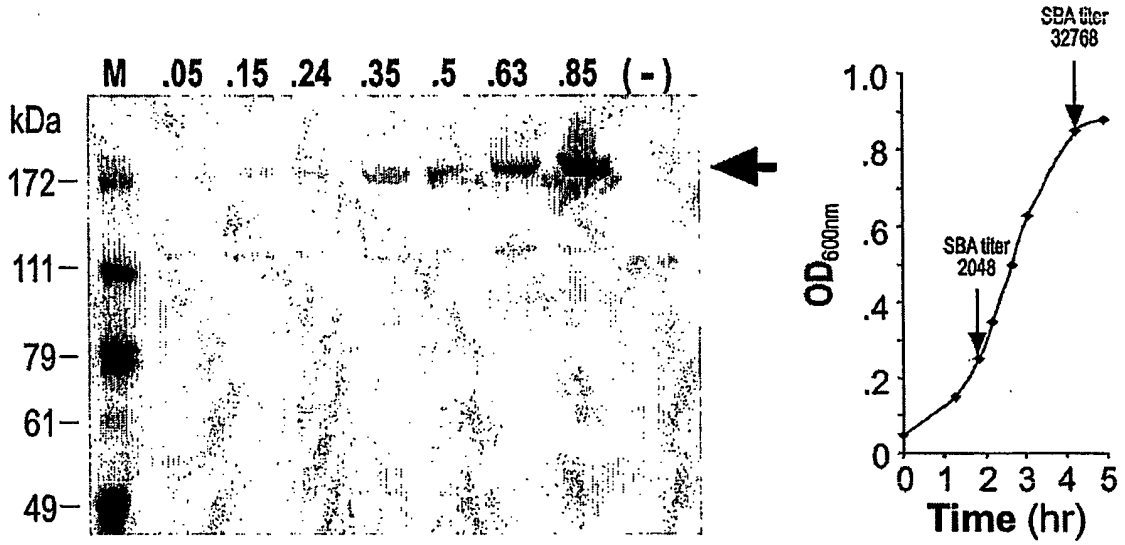


FIGURE 15B

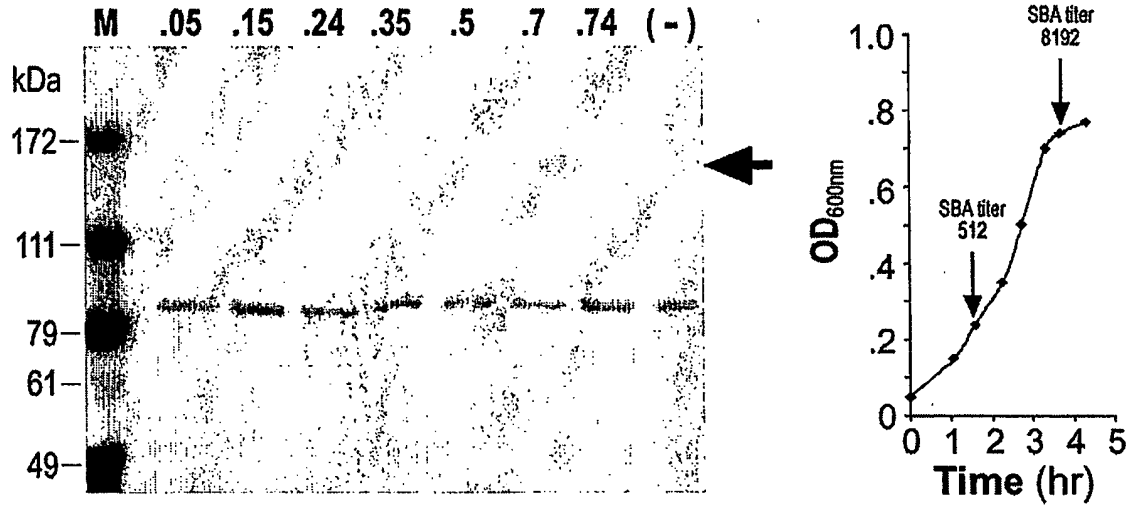


FIGURE 16

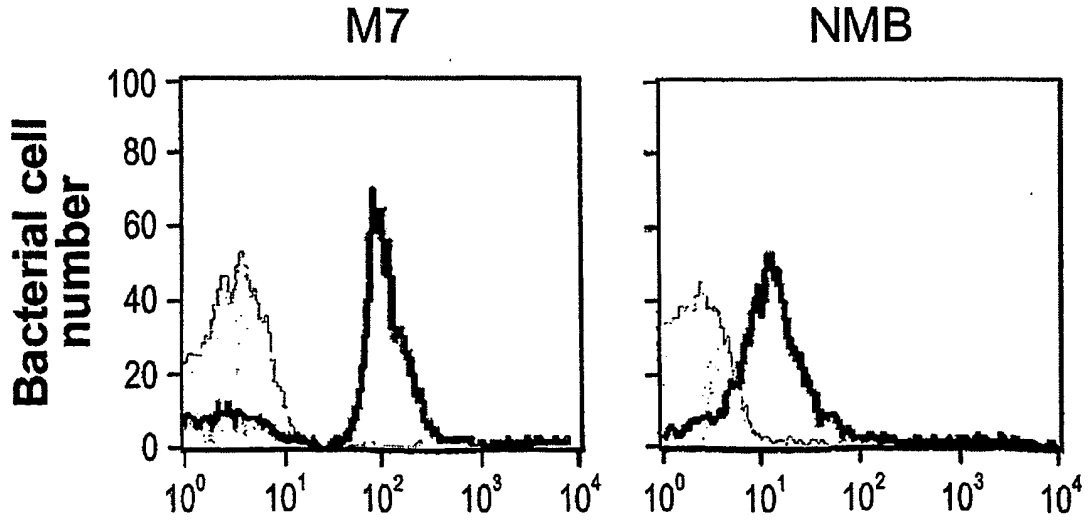


FIGURE 17

FIGURE 17A

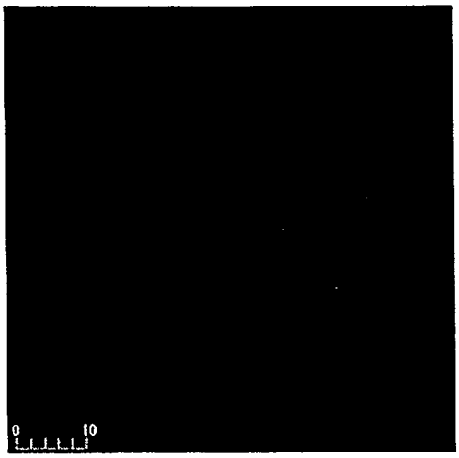


FIGURE 17B

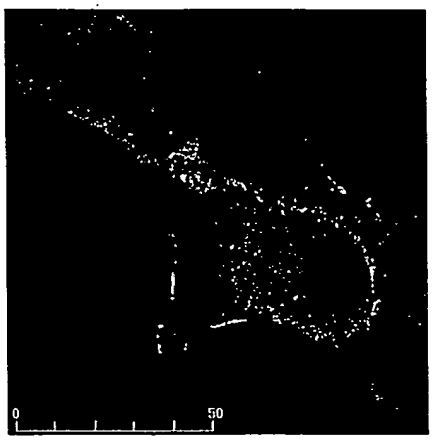


FIGURE 17C

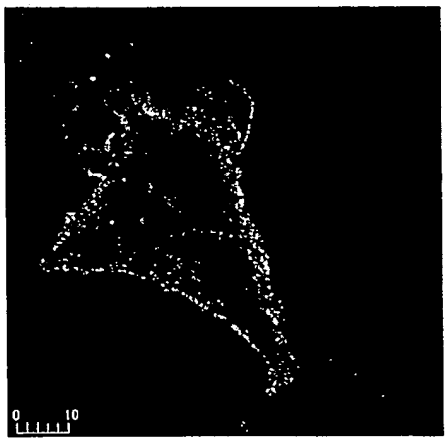


FIGURE 17D

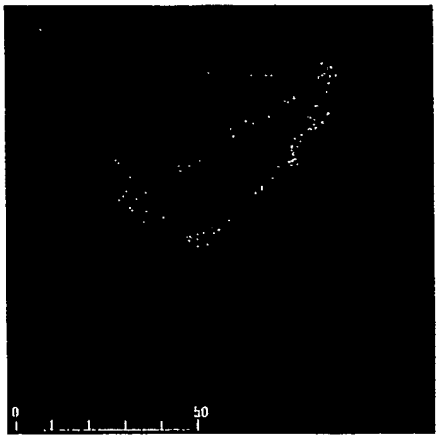


FIGURE 18

FIGURE 18A

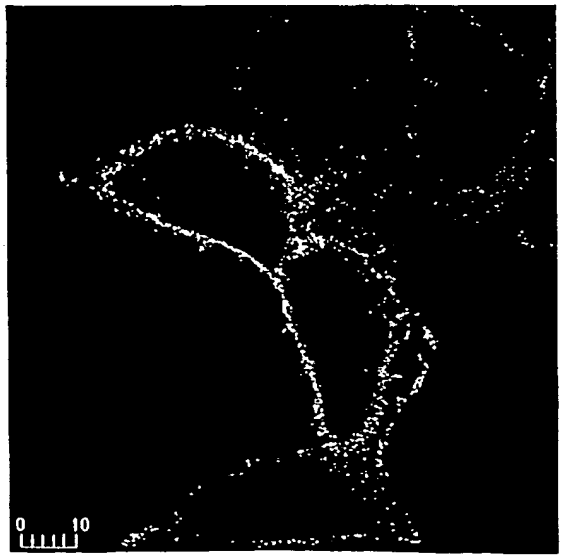


FIGURE 18B

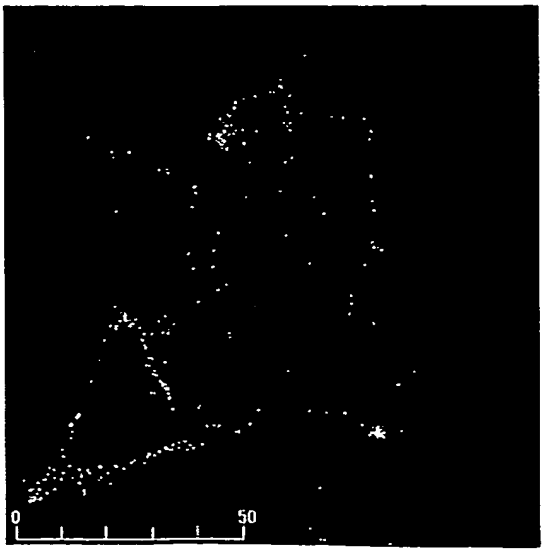


FIGURE 19

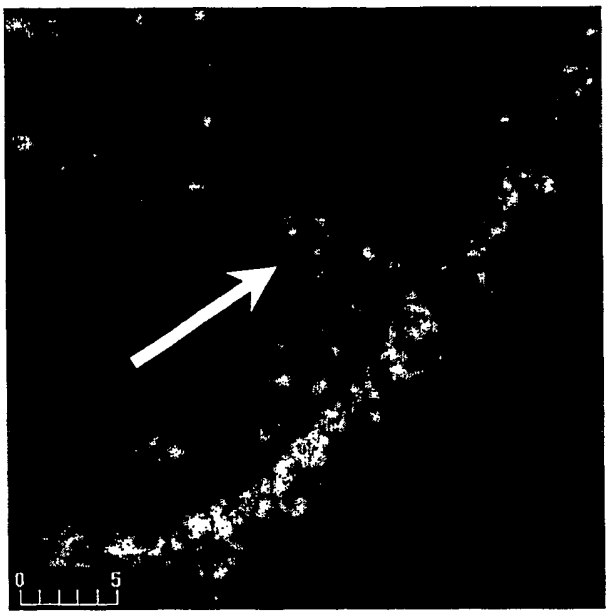


FIGURE 20

FIGURE 20A

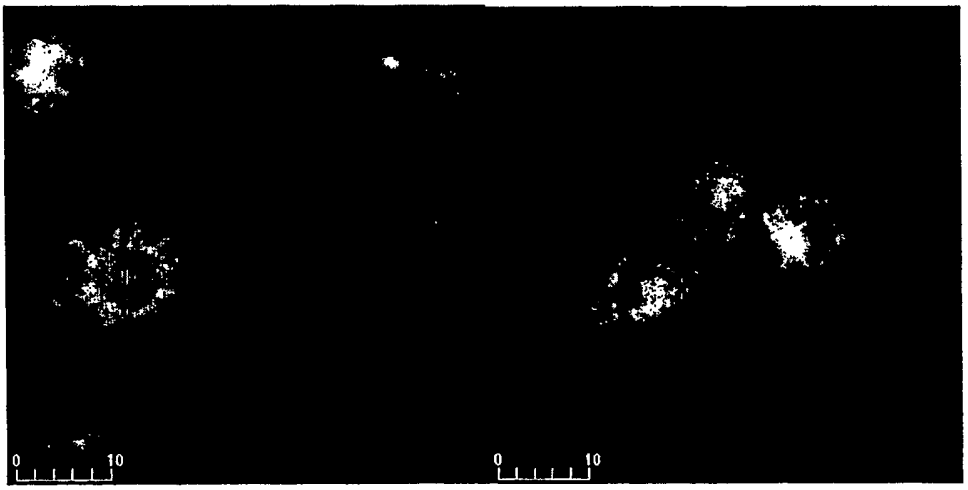


FIGURE 20B

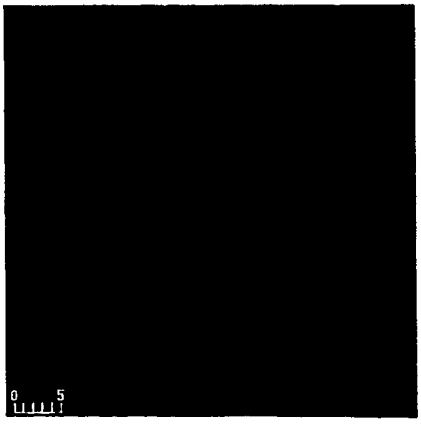


FIGURE 20C

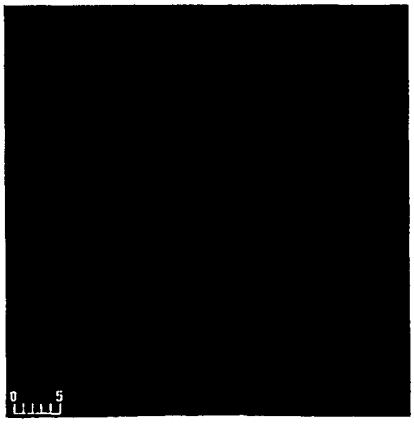


FIGURE 21

FIGURE 22

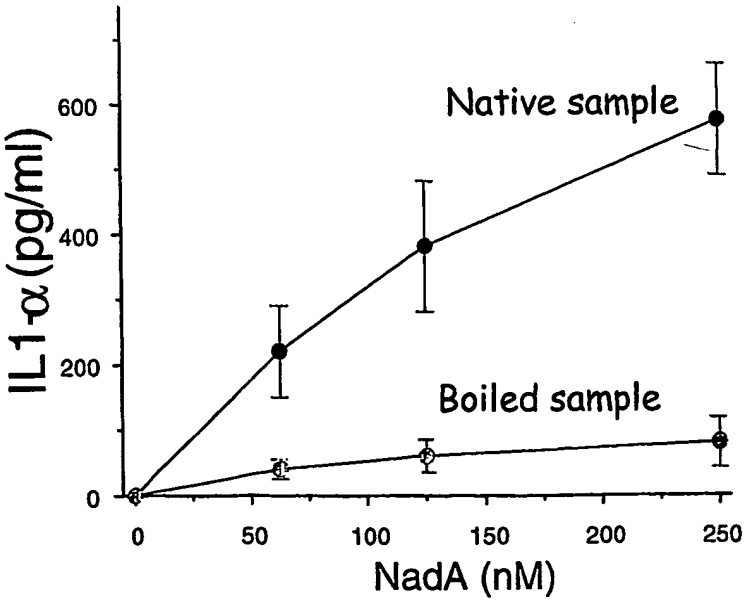


FIGURE 23

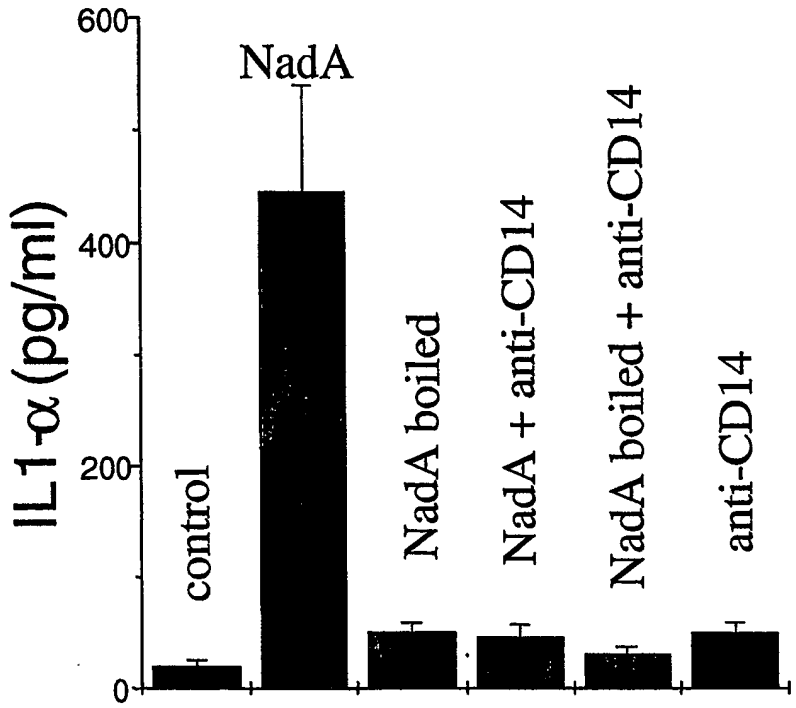


FIGURE 24

FIGURE 24A

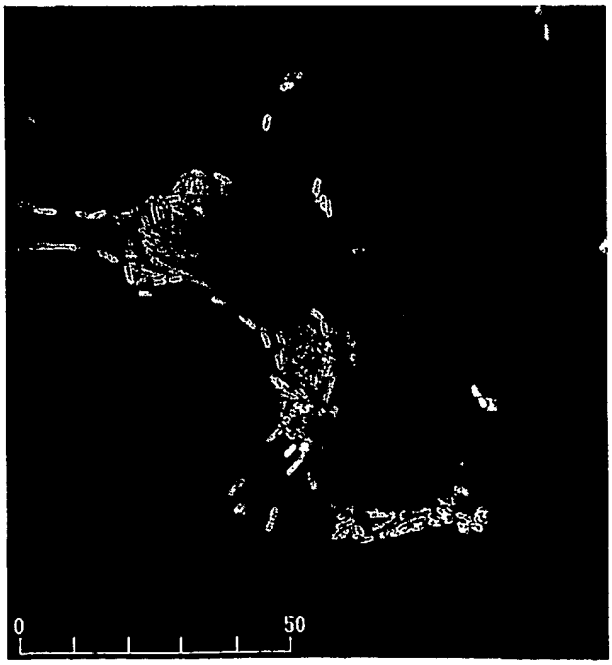


FIGURE 24B

FIGURE 25

FIGURE 25A

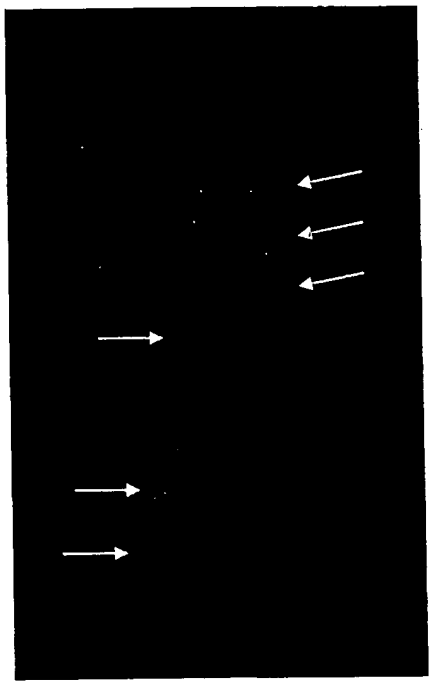


FIGURE 25B

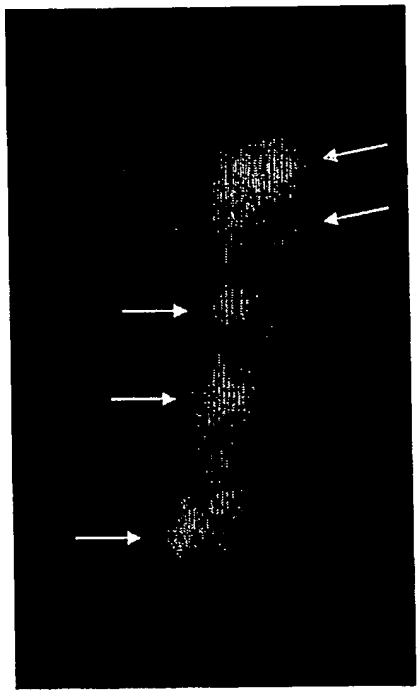


FIGURE 25C

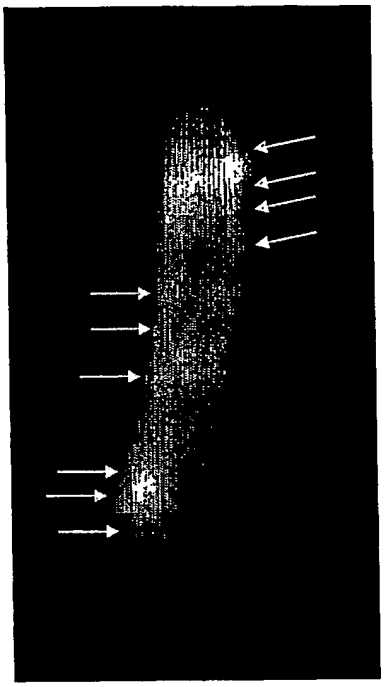


FIGURE 26

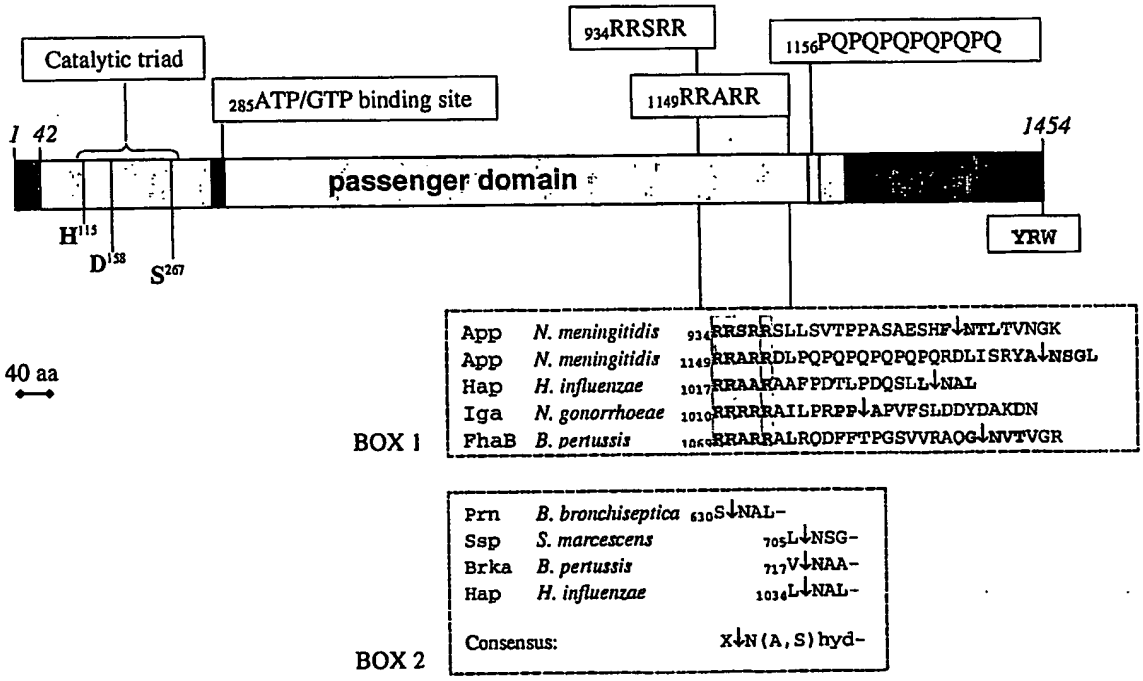


FIGURE 27

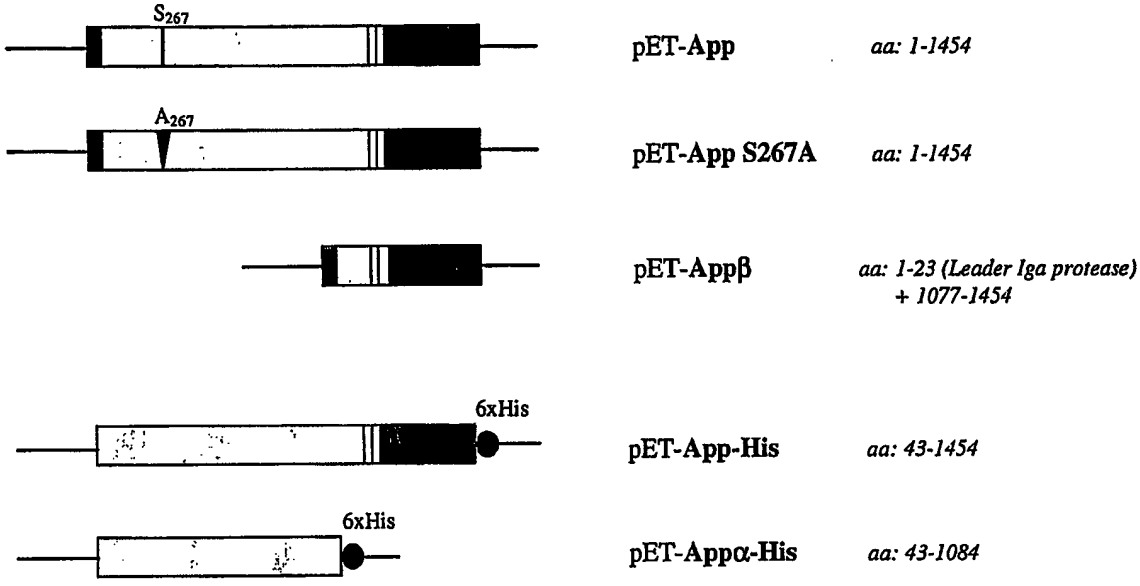


FIGURE 28

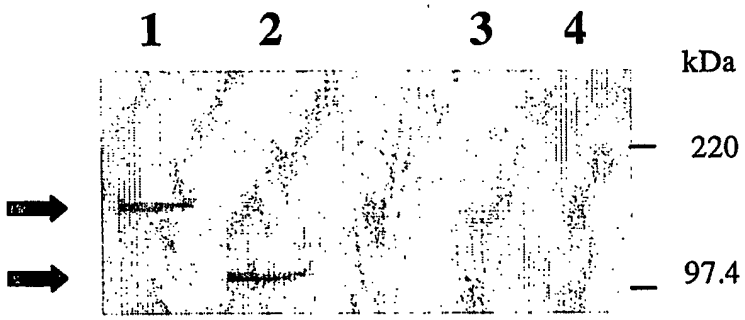


FIGURE 29

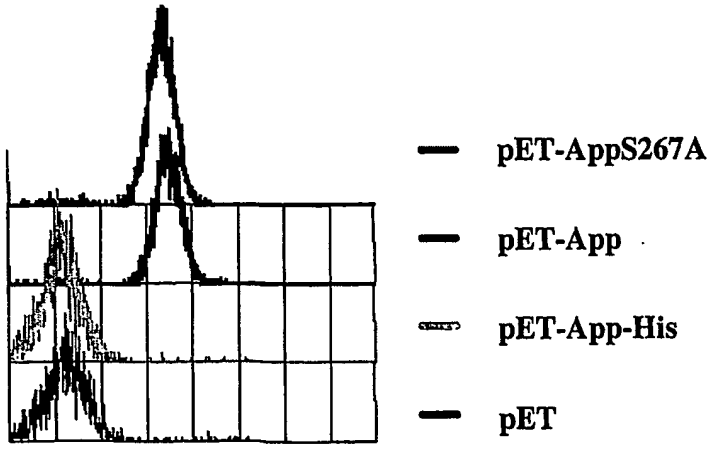


FIGURE 30

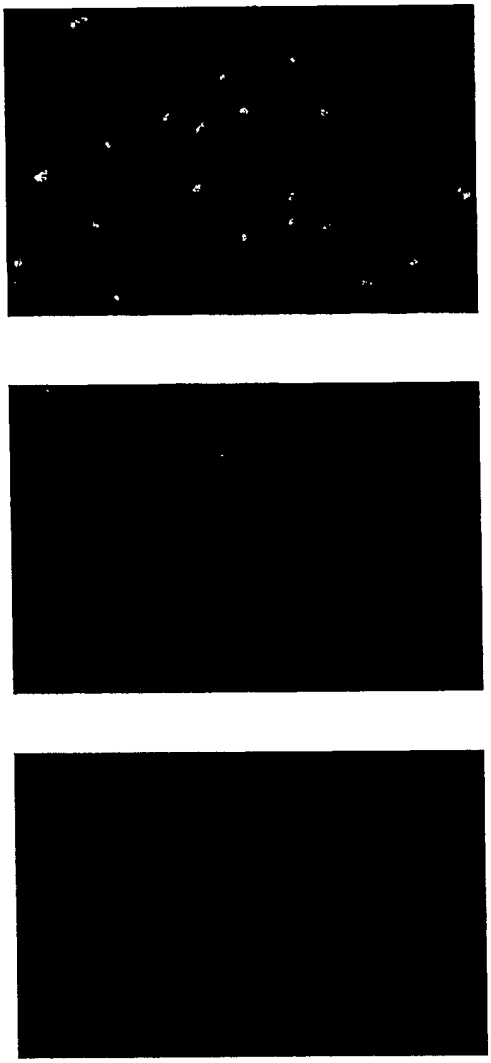


FIGURE 31

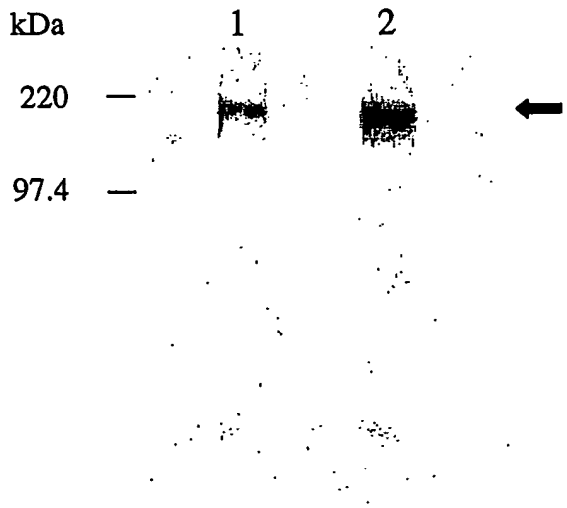


FIGURE 32

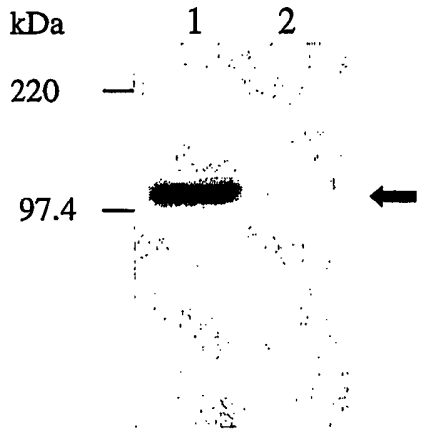


FIGURE 33

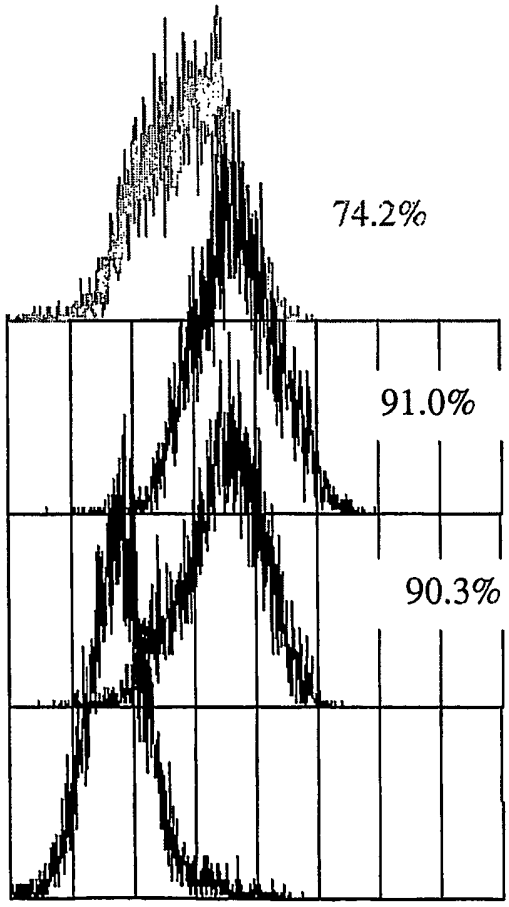


FIGURE 34

FIGURE 34A

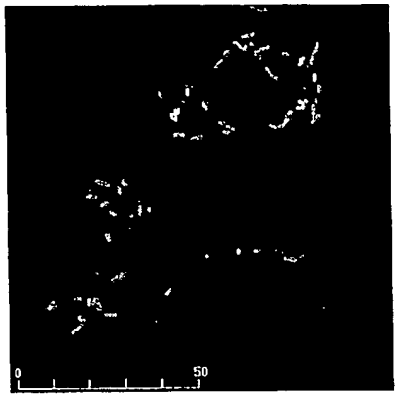


FIGURE 34B

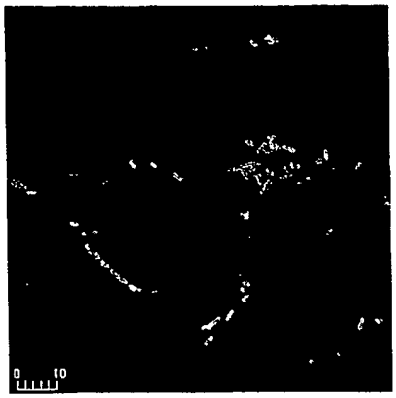


FIGURE 34C

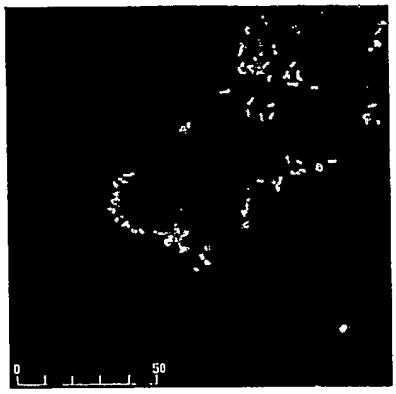


FIGURE 34D

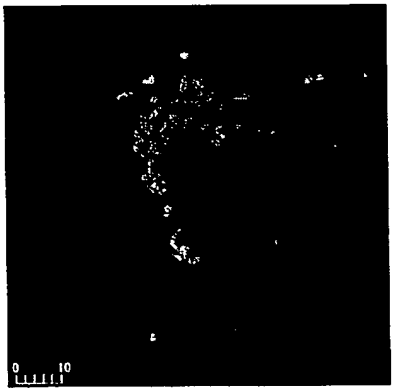


FIGURE 34E

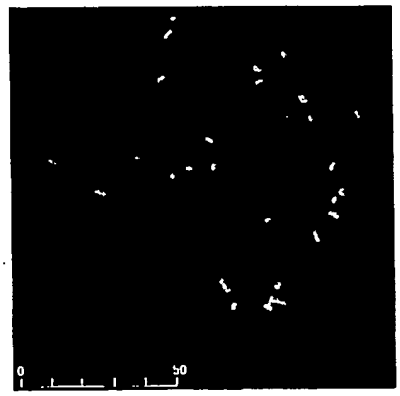


FIGURE 34F

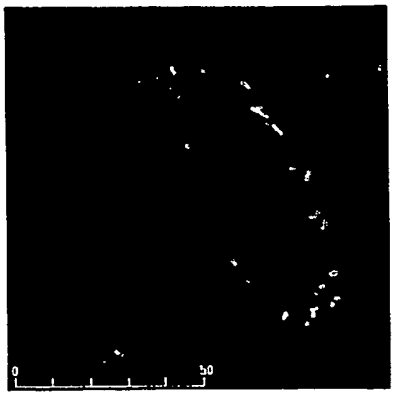


FIGURE 34G

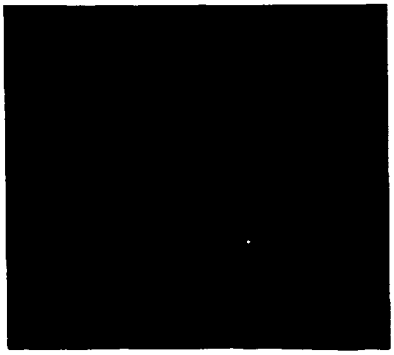


FIGURE 35

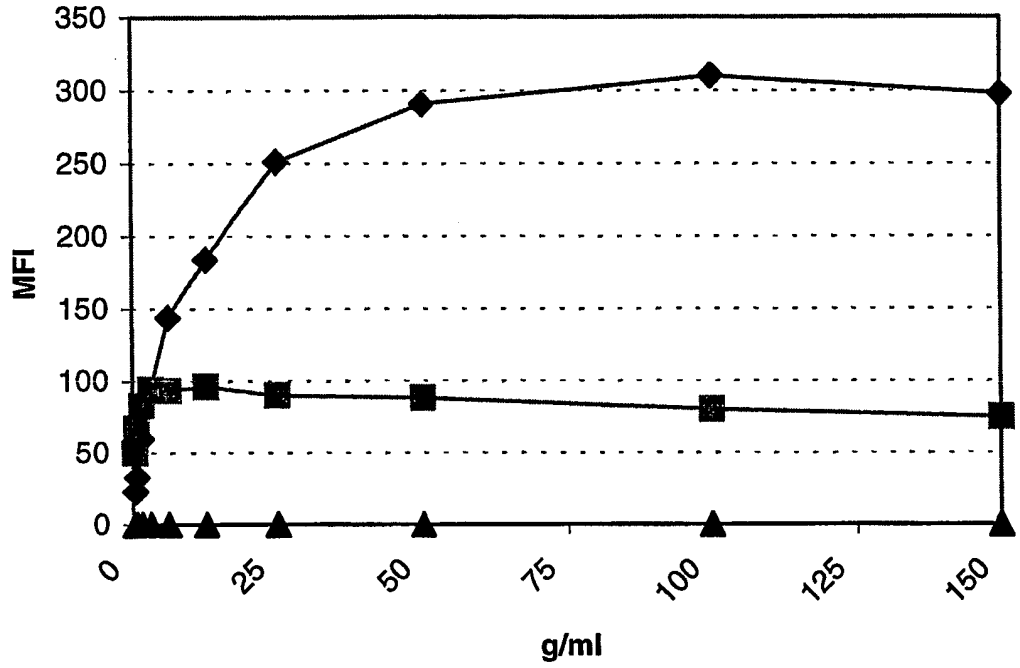


FIGURE 36

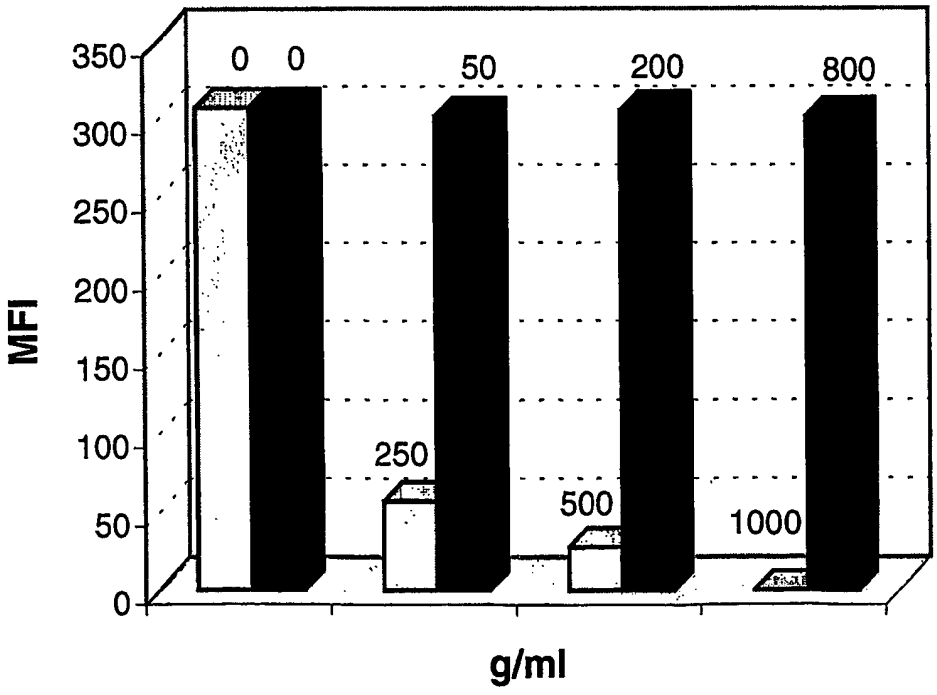


FIGURE 37

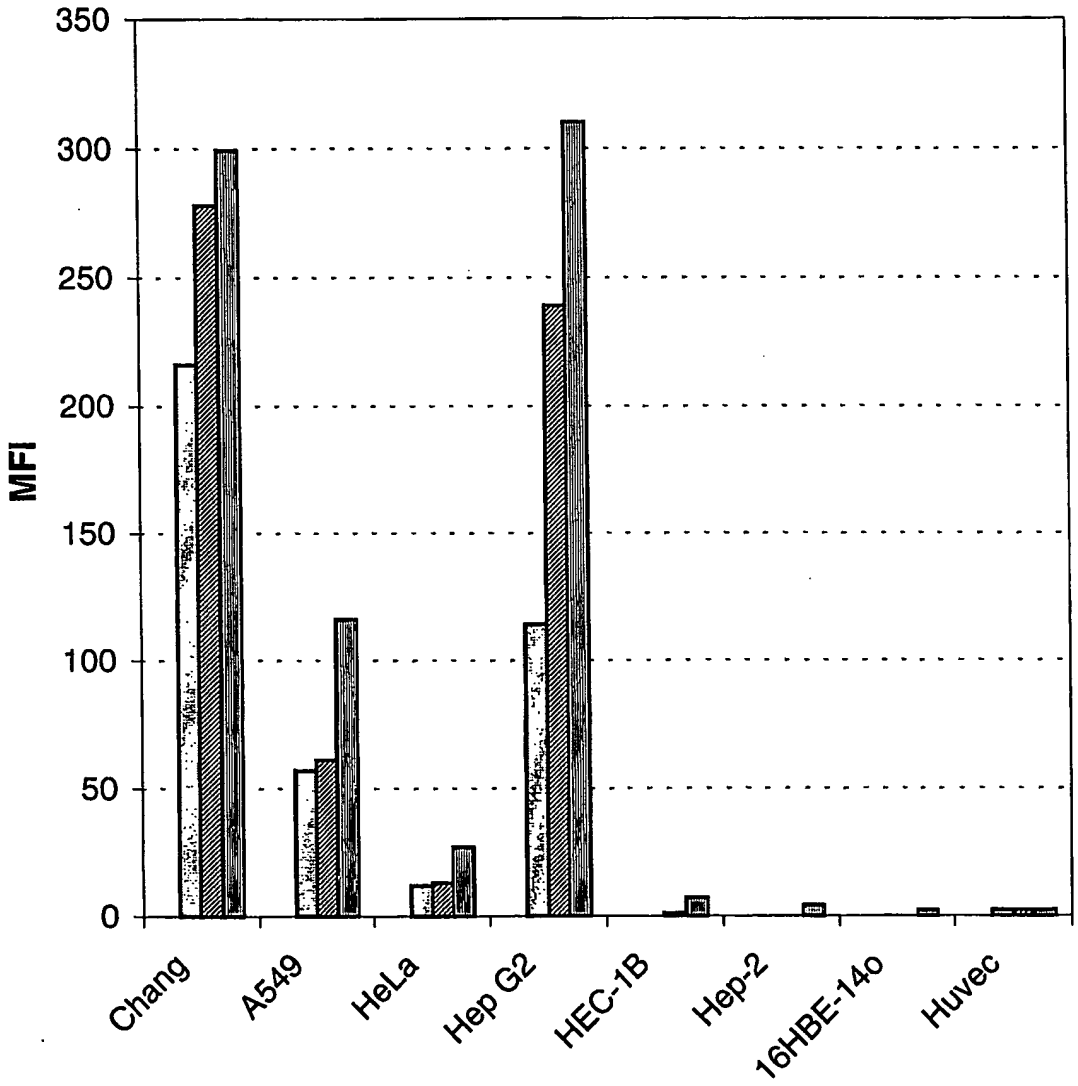

350
300
250
200
150
100
50
0
MFI
Chang
A549
HeLa
Hep G2
HEC-1B
Hep-2
16HBE-14o
Huvec

FIGURE 38

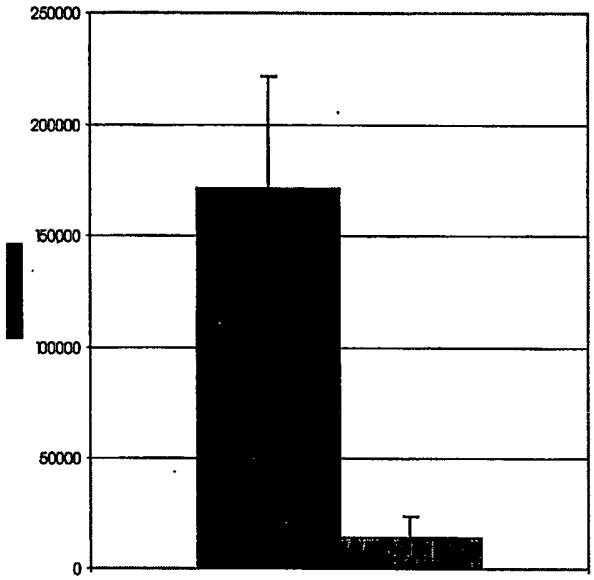


FIGURE 39

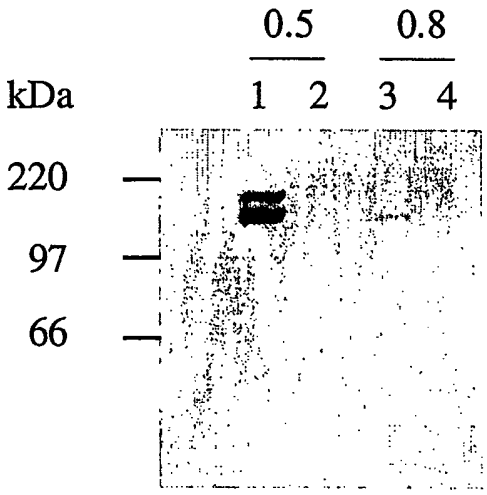
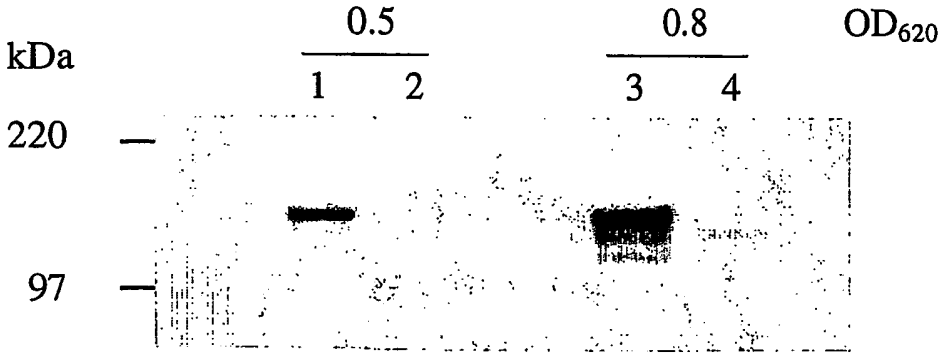


FIGURE 40



SEQUENCE LISTING

SEQ ID 1 – allele 1 of NadA

MKHFPSKVLTTAILATFCSGALAATSDDDVKKAATVAIVAAYNNGQEINGFKAGETIYDIGEDGTITQKDATAADVEADDFKGLGLKKVVTNLTKT
VNENKQNVDAKVKAAESEIEKLTTKLADTDAALADTDAALDETTNALNKLGENITTFAEETKTNIVKIDEKLEAVADTVDKHAEAFNDIADSLDET
NTKADEAVKTANEAKQTAEETKQNVDAKVKAAETAAGKAEAAAGTANTAADKAEAVAAKVTDIKADIATNKADIAKNSARIDSLDKNVANLRKETR
QGLAEQAALSGLFQPYNVGRFNVTAAVGGYKSESAVAIGTGFRFTENFAAKAGVAVGTSSGSSAAYHVGVNYEW

SEQ ID 2 – allele 2 of NadA

MKHFPSKVLTTAILATFCSGALAATNDDDVKKAATVAIAAAYNNGQEINGFKAGETIYDIDEDGTITKKDATAADVEADDFKGLGLKKVVTNLTKT
VNENKQNVDAKVKAAESEIEKLTTKLADTDAALDATTNALNKLGENITTFAEETKTNIVKIDEKLEAVADTVDKHAEAFNDIADSLDETNTKADEA
VKTANEAKQTAEETKQNVDAKVKAAETAAGKAEAAAGTANTAADKAEAVAAKVTDIKADIATNKDNIAKKANSADVYTREESDSKFVRIDGLNATT
EKLDTRLASAEKSITEHGTRLNGLDRTVSDLRKETRQGLAEQAALSGLFQPYNVGRFNVTAAVGGYKSESAVAIGTGFRFTENFAAKAGVAVGTSS
GSSAAYHVGVNYEW

SEQ ID 3 – allele 3 of NadA

MKHFPSKVLTTAILATFCSGALAATNDDDVKKAATVAIAAAYNNGQEINGFKAGETIYDIDEDGTITKKDATAADVEADDFKGLGLKKVVTNLTKT
VNENKQNVDAKVKAAESEIEKLTTKLADTDAALADTDAALDATTNALNKLGENITTFAEETKTNIVKIDEKLEAVADTVDKHAEAFNDIADSLDET
NTKADEAVKTANEAKQTAEETKQNVDAKVKAAETAAGKAEAAAGTANTAADKAEAVAAKVTDIKADIATNKDNIAKKANSADVYTREESDSKFVRI
DGLNATTEKLDTRLASAEKSIADHDTRLNGLDKTVSDLRKETRQGLAEQAALSGLFQPYNVGRFNVTAAVGGYKSESAVAIGTGFRFTENFAAKAG
VAVGTSSGSSAAYHVGVNYEW

SEQ ID 4 – allele 1 of NadA (first-ATG start)

MSMKHFPSKVLTTAILATFCSGALAATSDDDVKKAATVAIVAAYNNGQEINGFKAGETIYDIGEDGTITQKDATAADVEADDFKGLGLKKVVTNLT
KTVNENKQNVDAKVKAAESEIEKLTTKLADTDAALADTDAALDETTNALNKLGENITTFAEETKTNIVKIDEKLEAVADTVDKHAEAFNDIADSLD
ETNTKADEAVKTANEAKQTAEETKQNVDAKVKAAETAAGKAEAAAGTANTAADKAEAVAAKVTDIKADIATNKADIAKNSARIDSLDKNVANLRKE
TRQGLAEQAALSGLFQPYNVGRFNVTAAVGGYKSESAVAIGTGFRFTENFAAKAGVAVGTSSGSSAAYHVGVNYEW

SEQ ID 5 – allele 2 of NadA (first-ATG start)

MSMKHFPSKVLTTAILATFCSGALAATNDDDVKKAATVAIAAAYNNGQEINGFKAGETIYDIDEDGTITKKDATAADVEADDFKGLGLKKVVTNLT
KTVNENKQNVDAKVKAAESEIEKLTTKLADTDAALDATTNALNKLGENITTFAEETKTNIVKIDEKLEAVADTVDKHAEAFNDIADSLDETNTKAD
EAVKTANEAKQTAEETKQNVDAKVKAAETAAGKAEAAAGTANTAADKAEAVAAKVTDIKADIATNKDNIAKKANSADVYTREESDSKFVRIDGLNA
TTEKLDTRLASAEKSITEHGTRLNGLDRTVSDLRKETRQGLAEQAALSGLFQPYNVGRFNVTAAVGGYKSESAVAIGTGFRFTENFAAKAGVAVGT
SSGSSAAYHVGVNYEW

SEQ ID 6 – allele 3 of NadA (first-ATG start)

MSMKHFPSKVLTTAILATFCSGALAATNDDDVKKAATVAIAAAYNNGQEINGFKAGETIYDIDEDGTITKKDATAADVEADDFKGLGLKKVVTNLT
KTVNENKQNVDAKVKAAESEIEKLTTKLADTDAALADTDAALDATTNALNKLGENITTFAEETKTNIVKIDEKLEAVADTVDKHAEAFNDIADSLD
ETNTKADEAVKTANEAKQTAEETKQNVDAKVKAAETAAGKAEAAAGTANTAADKAEAVAAKVTDIKADIATNKDNIAKKANSADVYTREESDSKFV
RIDGLNATTEKLDTRLASAEKSIADHDTRLNGLDKTVSDLRKETRQGLAEQAALSGLFQPYNVGRFNVTAAVGGYKSESAVAIGTGFRFTENFAAK
AGVAVGTSSGSSAAYHVGVNYEW

SEQ ID 7 – variant allele 2 of NadA in strain ISS1024

MKHFPSKVLTTAILATFCSGALAATNDDDVKKAATVAIAAAYNNGQEINGFKAGETIYDIDEDGTITKKDATAADVEADDFKGLGLKKVVTNLTKT
VNENKQNVDAKVKAAESEIEKLTTKLADTDAALDATTNALNKLGENITTFAEETKTNIVKIDEKLEAVADTVDKHAEAFNDIADSLDETNTKADEA
VKTANEAKQTAEETKQNVDAKVKAAETAAGTANTAADKAEAVAAKVTDIKADIATNKDNIAKKANSADVYTREESDSKFVRIDGLNATTEKLDTRL
ASAEKSITEHGTRLNGLDRTVSDLRKETRQGLAEQAALSGLFQPYNVGRFNVTAAVGGYKSESAVAIGTGFRFTENFAAKAGVAVGTSSGSSAAYH
VGVNYEW

SEQ ID 8 – variant allele 2 of NadA (first-ATG start) in strain ISS1024

MSMKHFPSKVLTTAILATFCSGALAATNDDDVKKAATVAIAAAYNNGQEINGFKAGETIYDIDEDGTITKKDATAADVEADDFKGLGLKKVVTNLT
KTVNENKQNVDAKVKAAESEIEKLTTKLADTDAALDATTNALNKLGENITTFAEETKTNIVKIDEKLEAVADTVDKHAEAFNDIADSLDETNTKAD
EAVKTANEAKQTAEETKQNVDAKVKAAETAAGTANTAADKAEAVAAKVTDIKADIATNKDNIAKKANSADVYTREESDSKFVRIDGLNATTEKLDT
RLASAEKSITEHGTRLNGLDRTVSDLRKETRQGLAEQAALSGLFQPYNVGRFNVTAAVGGYKSESAVAIGTGFRFTENFAAKAGVAVGTSSGSSAA
YHVGVNYEW

SEQ ID 9 – variant allele 3 of NadA in strains 973-1720 and ISS759

MQHFPSKVLTTAILATFCSGALAATNDDDVKKAATVAIAAAYNNGQEINGFKAGETIYDIDEDGTITKKDATAADVEADDFKGLGLKKVVTNLTKT
VNENKQNVDAKVKAAESEIEKLTTKLADTDAALADTDAALDATTNALNKLGENITTFAEETKTNIVKIDEKLEAVADTVDKHAEAFNDIADSLDET
NTKADEAVKTANEAKQTAEETKQNVDAKVKAAETAAGKAEAAAGTANTAADKAEAVAAKVTDIKADIATNKDNIAKKANSADVYTREESDSKFVRI
DGLNATTEKLDTRLASAEKSIADHDTRLNGLDKTVSDLRKETRQGLAEQAALSGLFQPYNVGRFNVTAAVGGYKSESAVAIGTGFRFTENFAAKAG
VAVGTSSGSSAAYHVGVNYEW

SEQ ID 10 – variant allele 3 of NadA (first-ATG start) in strains 973-1720 and ISS759

MSMQHFPSKVLTTAILATFCSGALAATNDDDVKKAATVAIAAAYNNGQEINGFKAGETIYDIDEDGTITKKDATAADVEADDFKGLGLKKVVTNLT
KTVNENKQNVDAKVKAAESEIEKLTTKLADTDAALADTDAALDATTNALNKLGENITTFAEETKTNIVKIDEKLEAVADTVDKHAEAFNDIADSLD
ETNTKADEAVKTANEAKQTAEETKQNVDAKVKAAETAAGKAEAAAGTANTAADKAEAVAAKVTDIKADIATNKDNIAKKANSADVYTREESDSKFV
RIDGLNATTEKLDTRLASAEKSIADHDTRLNGLDKTVSDLRKETRQGLAEQAALSGLFQPYNVGRFNVTAAVGGYKSESAVAIGTGFRFTENFAAK
AGVAVGTSSGSSAAYHVGVNYEW

SEQ ID 11 – NadA allele 1/2 chimera (strain 95330)

MKHFPSKVLTTAILATFCSGALAATNDDDVKKAATVAIAAAYNNGQEINGFKAGETIYDIDEDGTITKKDATAADVEADDFKGLGLKKVVTNLTKT
VNENKQNVDAKVKAAESEIEKLTTKLADTDAALDATTNALNKLGENITTFAEETKTNIVKIDEKLEAVADTVDKHAEAFNDIADSLDETNTKADEA
VKTANEAKQTAEETKQNVDAKVKAAETAAGKAEAAAGTANTAADKAEAVAAKVTDIKADIATNKADIAKNSARIDSLDKNVANLRKETRQGLAEQA
ALSGLFQPYNVGRFNVTAAVGGYKSESAVAIGTGFRFTENFAAKAGVAVGTSSGSSAAYHVGVNYEW

SEQ ID 12 – NadA allele 1/2 chimera (strain 95330) (first-ATG start)

MSMKHFPSKVLTTAILATFCSGALAATNDDDVKKAATVAIAAAYNNGQEINGFKAGETIYDIDEDGTITKKDATAADVEADDFKGLGLKKVVTNLT
KTVNENKQNVDAKVKAAESEIEKLTTKLADTDAALDATTNALNKLGENITTFAEETKTNIVKIDEKLEAVADTVDKHAEAFNDIADSLDETNTKAD
EAVKTANEAKQTAEETKQNVDAKVKAAETAAGKAEAAAGTANTAADKAEAVAAKVTDIKADIATNKADIAKNSARIDSLDKNVANLRKETRQGLAE
QAALSGLFQPYNVGRFNVTAAVGGYKSESAVAIGTGFRFTENFAAKAGVAVGTSSGSSAAYHVGVNYEW

SEQ ID 13 – NadA allele C

MKHFPSKVLTAAILAALSGSAMADNAPTADEIAKAALVNSYNNTQDINGFTVGDTIYDIKNDKITKKEATEADVEADDFKGLGLKEVVAQHDQSLA
DLTETVNENSEALVKTAAVVNDISADVKANTAAIGENKAAIATKADKTELDKVSGKVTENETAIGKKANSADVYTKAEVYTKQESDNRFVKISDGI
GNLNTTANGLETRLAAAEQSVADHGTRLASAEKSITEHGTRLNGLDRTVSDLRKETRQGLAEQAALSGLFQPYNVGRFNVTAAVGGYKSESAVAIG
TGFRFTENFAAKAGVAVGTSSGSSAAYHVGVNYEW

SEQ ID 14 – NadA allele C (first-ATG start)

MSMKHFPSKVLTAAILAALSGSAMADNAPTADEIAKAALVNSYNNTQDINGFTVGDTIYDIKNDKITKKEATEADVEADDFKGLGLKEVVAQHDQS
LADLTETVNENSEALVKTAAVVNDISADVKANTAAIGENKAAIATKADKTELDKVSGKVTENETAIGKKANSADVYTKAEVYTKQESDNRFVKISD
GIGNLNTTANGLETRLAAAEQSVADHGTRLASAEKSITEHGTRLNGLDRTVSDLRKETRQGLAEQAALSGLFQPYNVGRFNVTAAVGGYKSESAVA
IGTGFRFTENFAAKAGVAVGTSSGSSAAYHVGVNYEW

SEQ ID 15 – coding sequence for SEQ ID 13

ATGAAACACTTTCCATCCAAAGTACTGACCGCAGCCATCCTTGCCGCCCTCAGCGGCAGCGCAATGGCAGACAACGCCCCCACCGCTGACGAAATT
GCCAAAGCCGCCCTAGTTAACTCCTACAACAATACCCAAGACATCAACGGATTCACAGTCGGAGACACCATCTACGACATTAAAAATGACAAGATT
ACCAAAAAAGAAGCAACAGAAGCCGATGTTGAAGCTGACGACTTTAAAGGTCTGGGTCTGAAAGAAGTCGTGGCTCAACACGACCAAAGCCTTGCC
GACCTGACCGAAACCGTCAATGAAAACAGCGAAGCATTGGTAAAAACCGCCGCGGTTGTCAATGACATCAGTGCCGATGTCAAAGCCAACACAGCA
GCTATCGGGGAAAACAAAGCTGCTATCGCTACAAAAGCAGACAAAACCGAACTGGATAAAGTGTCCGGCAAAGTAACCGAGAACGAGACTGCTATC
GGTAAAAAAGCAAACAGTGCCGACGTGTACACTAAAGCTGAGGTGTACACCAAACAAGAGTCTGACAACAGATTTGTCAAAATTAGTGACGGAATC
GGTAATCTGAACACTACTGCCAATGGATTGGAGACACGCTTGGCCGCTGCCGAACAATCCGTTGCAGACCACGGTACGCGCTTGGCTTCTGCCGAA
AAATCCATTACCGAACACGGTACGCGCCTGAACGGTTTGGATAGAACAGTGTCAGACCTGCGTAAAGAAACCCGCCAAGGCCTTGCAGAACAAGCC
GCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTCGGTTCAATGTAACGGCTGCAGTCGGCGGCTACAAATCCGAATCGGCAGTCGCCATCGGT
ACCGGCTTCCGCTTTACCGAAAACTTTGCCGCCAAAGCAGGCGTGGCAGTCGGCACTTCGTCCGGTTCTTCCGCAGCCTACCATGTCGGCGTCAAT
TACGAGTGGTAA

SEQ ID 16 – forward primer

GTCGACGTCCTCGATTACGAAG

SEQ ID 17 – reverse primer

CGAGGCGATTGTCAAACCGTTC

SEQ ID 18 – forward primer
cgcggatccgctagcGGACACACTTATTTCGG

SEQ ID 19 – reverse primer
cccgctcgagCCAGCGGTAGCCTAATTTG

SEQ ID 20 – forward primer
cgcggatccgctagcAAAACAACCGACAAACGG

SEQ ID 21 – reverse primer
cccgctcgagTTACCAGCGGTAGCCTAATTTG

SEQ ID 22 – mutagenesis primer
CTCATTTGGCGACgctGGCTCACCAATGTTTATCTATGATG

SEQ ID 23 – mutagenesis primer
CATCATAGATAAACATTGGTGAGCCagcGTCGCCAAATGAG

SEQ ID 24 – forward primer
cgcggatccgctagcGGACACACTTATTTCGG

SEQ ID 25 – reverse primer
cccgctcgagCAGCGCGTCAAGGCTT

SEQ ID 26 – forward primer
gggaattccatatgaaagccaaacgttttaaaattaacgccatatccttatccatctttcttgcctatgcccttacgccatactcagaagcggcta
gcGACAACGCGCAAAGCCTTGACGCGCT

SEQ ID 27 – reverse primer
cccgctcgagTTACCAGCGGTAGCCTAATTTG

SEQ ID 28 – knockout primer
gctctagaggaggctgtcgaaacc

SEQ ID 29 – knockout primer
tcccccgggcggttgtccgtttgtcg

SEQ ID 30 – knockout primer
tcccccggggcgggcatcaaattaggc

SEQ ID 31 – knockout primer
Cccgctcgagcgcaaccgtccgctgac

SEQ ID 32 – SEQ ID 650 from WO99/24578
MKTTDKRTTETHRKAPKTGRIRFSPAYLAICLSFGILPQAWAGHTYFGINYQYYRDFAENKGKFAVGAKDIEVYNKKGELVGKSMTKAPMIDFSVV
SRNGVAALVGDQYIVSVAHNGGYNNVDFGAEGRNPDQHRFTYKIVKRNNYKAGTKGHPYGGDYHMPRLHKFVTDAEPVEMTSYMDGRKYIDQNNYP
DRVRIGAGRQYWRSDEDEPNNRESSYHIASAYSWLVGGNTFAQNGSGGGTVNLGSEKIKHSPYGFLPTGGSFGDSGSPMFIYDAQKQKWLINGVLQ
TGNPYIGKSNGFQLVRKDWFYDEIFAGDTHSVFYEPRQNGKYSFNDDNNGTGKINAKHEHNSLPNRLKTRTVQLFNVSLSETAREPVYHAAGGVNS
YRPRLNNGENISFIDEGKGELILTSNINQGAGGLYFQGDFTVSPENNETWQGAGVHISEDSTVTWKVNGVANDRLSKIGKGTLHVQAKGENQGSIS
VGDGTVILDQQADDKGKKQAFSEIGLVSGRGTVQLNADNQFNPDKLYFGFRGGRLDLNGHSLSFHRIQNTDEGAMIVNHNQDKESTVTITGNKDIA
TTGNNNSLDSKKEIAYNGWFGEKDTTKTNGRLNLVYQPAAEDRTLLLSGGTNLNGNITQTNGKLFFSGRPTPHAYNHLNDHWSQKEGIPRGEIVWD
NDWINRTFKAENFQIKGGQAVVSRNVAKVKGDWHLSNHAQAVFGVAPHQSHTICTRSDWTGLTNCVEKTITDDKVIASLTKTDISGNVDLADHAHL
NLTGLATLNGNLSANGDTRYTVSHNATQNGNLSLVGNAQATFNQATLNGNTSASGNASFNLSDHAVQNGSLTLSGNAKANVSHSALNGNVSLADKA
VFHFESSRFTGQISGGKDTALHLKDSEWTLPSGTELGNLNLDNATITLNSAYRHDAAGAQTGSATDAPRRRSRRSRRSLLSVTPPTSVESRFNTLT
VNGKLNGQGTFRFMSELFGYRSDKLKLAESSEGTYTLAVNNTGNEPASLEQLTVVEGKDNKPLSENLNFTLQNEHVDAGAWRYQLIRKDGEFRLHN
PVKEQELSDKLGKAEAKKQAEKDNAQSLDALIAAGRDAVEKTESVAEPARQAGGENVGIMQAEEEKKRVQADKDTALAKQREAETRPATTAFPRAR
RARRDLPQLQPQPQPQPQPQRDLISRYANSGLSEFSATLNSVFAVQDELDRVFAEDRRNAVWTSGIRDTKHYRSQDFRAYRQQTDLRQIGMQKNLGSG
RVGILFSHNRTENTFDDGIGNSARLAHGAVFGQYGIDRFYIGISAGAGFSSGSLSDGIGGKIRRRVLHYGIQARYRAGFGGFGIEPHIGATRYFVQ
KADYRYENVNIATPGLAFNRYRAGIKADYSFKPAQHISITPYLSLSYTDAASGKVRTRVNTAVLAQDFGKTRSAEWGVNAEIKGFTLSLHAAAAKG
PQLEAQHSAGIKLGYRW

SEQ ID 33 – App domain derivative

MKTTDKRTTETHRKAPKTGRIRFSPAYLAICLSFGILPQAWAGHTYFGINYQYYRDFAENKGKFAVGAKDIEVYNKKGELVGKSMTKAPMIDFSVV
SRNGVAALVGDQYIVSVAHNGGYNNVDFGAEGRNPDQHRFTYKIVKRNNYKAGTKGHPYGGDYHMPRLHKFVTDAEPVEMTSYMDGRKYIDQNNYP
DRVRIGAGRQYWRSDEDEPNNRESSYHIASAYSWLVGGNTFAQNGSGGGTVNLGSEKIKHSPYGFLPTGGSFGDSGSPMFIYDAQKQKWLINGVLQ
TGNPYIGKSNGFQLVRKDWFYDEIFAGDTHSVFYEPRQNGKYSFNDDNNGTGKINAKHEHNSLPNRLKTRTVQLFNVSLSETAREPVYHAAGGVNS
YRPRLNNGENISFIDEGKGELILTSNINQGAGGLYFQGDFTVSPENNETWQGAGVHISEDSTVTWKVNGVANDRLSKIGKGTLHVQAKGENQGSIS
VGDGTVILDQQADDKGKKQAFSEIGLVSGRGTVQLNADNQFNPDKLYFGFRGGRLDLNGHSLSFHRIQNTDEGAMIVNHNQDKESTVTITGNKDIA
TTGNNNSLDSKKEIAYNGWFGEKDTTKTNGRLNLVYQPAAEDRTLLLSGGTNLNGNITQTNGKLFFSGRPTPHAYNHLNDHWSQKEGIPRGEIVWD
NDWINRTFKAENFQIKGGQAVVSRNVAKVKGDWHLSNHAQAVFGVAPHQSHTICTRSDWTGLTNCVEKTITDDKVIASLTKTDISGNVDLADHAHL
NLTGLATLNGNLSANGDTRYTVSHNATQNGNLSLVGNAQATFNQATLNGNTSASGNASFNLSDHAVQNGSLTLSGNAKANVSHSALNGNVSLADKA
VFHFESSRFTGQISGGKDTALHLKDSEWTLPSGTELGNLNLDNATITLNSAYRHDAAGAQTGSATDAPRRRSRRSRRSLLSVTPPTSVESRF

SEQ ID 34 – App domain derivative

MKTTDKRTTETHRKAPKTGRIRFSPAYLAICLSFGILPQAWAGHTYFGINYQYYRDFAENKGKFAVGAKDIEVYNKKGELVGKSMTKAPMIDFSVV
SRNGVAALVGDQYIVSVAHNGGYNNVDFGAEGRNPDQHRFTYKIVKRNNYKAGTKGHPYGGDYHMPRLHKFVTDAEPVEMTSYMDGRKYIDQNNYP
DRVRIGAGRQYWRSDEDEPNNRESSYHIASAYSWLVGGNTFAQNGSGGGTVNLGSEKIKHSPYGFLPTGGSFGDSGSPMFIYDAQKQKWLINGVLQ
TGNPYIGKSNGFQLVRKDWFYDEIFAGDTHSVFYEPRQNGKYSFNDDNNGTGKINAKHEHNSLPNRLKTRTVQLFNVSLSETAREPVYHAAGGVNS
YRPRLNNGENISFIDEGKGELILTSNINQGAGGLYFQGDFTVSPENNETWQGAGVHISEDSTVTWKVNGVANDRLSKIGKGTLHVQAKGENQGSIS
VGDGTVILDQQADDKGKKQAFSEIGLVSGRGTVQLNADNQFNPDKLYFGFRGGRLDLNGHSLSFHRIQNTDEGAMIVNHNQDKESTVTITGNKDIA
TTGNNNSLDSKKEIAYNGWFGEKDTTKTNGRLNLVYQPAAEDRTLLLSGGTNLNGNITQTNGKLFFSGRPTPHAYNHLNDHWSQKEGIPRGEIVWD
NDWINRTFKAENFQIKGGQAVVSRNVAKVKGDWHLSNHAQAVFGVAPHQSHTICTRSDWTGLTNCVEKTITDDKVIASLTKTDISGNVDLADHAHL
NLTGLATLNGNLSANGDTRYTVSHNATQNGNLSLVGNAQATFNQATLNGNTSASGNASFNLSDHAVQNGSLTLSGNAKANVSHSALNGNVSLADKA
VFHFESSRFTGQISGGKDTALHLKDSEWTLPSGTELGNLNLDNATITLNSAYRHDAAGAQTGSATDAPRRRSRRSRRSLLSVTPPTSVESRFNTLT
VNGKLNGQGTFRFMSELFGYRSDKLKLAESSEGTYTLAVNNTGNEPASLEQLTVVEGKDNKPLSENLNFTLQNEHVDAGAWRYQLIRKDGEFRLHN
PVKEQELSDKLGKAEAKKQAEKDNAQSLDALIAAGRDAVEKTESVAEPARQAGGENVGIMQAEEEKKRVQADKDTALAKQREAETRPATTAFPRAR
RARRDLPQLQPQPQPQPQRDLISRYA

SEQ ID 35 – App domain derivative

GHTYFGINYQYYRDFAENKGKFAVGAKDIEVYNKKGELVGKSMTKAPMIDFSVVSRNGVAALVGDQYIVSVAHNGGYNNVDFGAEGRNPDQHRFTY
KIVKRNNYKAGTKGHPYGGDYHMPRLHKFVTDAEPVEMTSYMDGRKYIDQNNYPDRVRIGAGRQYWRSDEDEPNNRESSYHIASAYSWLVGGNTFA
QNGSGGGTVNLGSEKIKHSPYGFLPTGGSFGDSGSPMFIYDAQKQKWLINGVLQTGNPYIGKSNGFQLVRKDWFYDEIFAGDTHSVFYEPRQNGKY
SFNDDNNGTGKINAKHEHNSLPNRLKTRTVQLFNVSLSETAREPVYHAAGGVNSYRPRLNNGENISFIDEGKGELILTSNINQGAGGLYFQGDFTV
SPENNETWQGAGVHISEDSTVTWKVNGVANDRLSKIGKGTLHVQAKGENQGSISVGDGTVILDQQADDKGKKQAFSEIGLVSGRGTVQLNADNQFN
PDKLYFGFRGGRLDLNGHSLSFHRIQNTDEGAMIVNHNQDKESTVTITGNKDIATTGNNNSLDSKKEIAYNGWFGEKDTTKTNGRLNLVYQPAAED
RTLLLSGGTNLNGNITQTNGKLFFSGRPTPHAYNHLNDHWSQKEGIPRGEIVWDNDWINRTFKAENFQIKGGQAVVSRNVAKVKGDWHLSNHAQAV
FGVAPHQSHTICTRSDWTGLTNCVEKTITDDKVIASLTKTDISGNVDLADHAHLNLTGLATLNGNLSANGDTRYTVSHNATQNGNLSLVGNAQATF
NQATLNGNTSASGNASFNLSDHAVQNGSLTLSGNAKANVSHSALNGNVSLADKAVFHFESSRFTGQISGGKDTALHLKDSEWTLPSGTELGNLNLD
NATITLNSAYRHDAAGAQTGSATDAPRRRSRRSRRSLLSVTPPTSVESRF

SEQ ID 36 – App domain derivative

GHTYFGINYQYYRDFAENKGKFAVGAKDIEVYNKKGELVGKSMTKAPMIDFSVVSRNGVAALVGDQYIVSVAHNGGYNNVDFGAEGRNPDQHRFTY
KIVKRNNYKAGTKGHPYGGDYHMPRLHKFVTDAEPVEMTSYMDGRKYIDQNNYPDRVRIGAGRQYWRSDEDEPNNRESSYHIASAYSWLVGGNTFA
QNGSGGGTVNLGSEKIKHSPYGFLPTGGSFGDSGSPMFIYDAQKQKWLINGVLQTGNPYIGKSNGFQLVRKDWFYDEIFAGDTHSVFYEPRQNGKY
SFNDDNNGTGKINAKHEHNSLPNRLKTRTVQLFNVSLSETAREPVYHAAGGVNSYRPRLNNGENISFIDEGKGELILTSNINQGAGGLYFQGDFTV
SPENNETWQGAGVHISEDSTVTWKVNGVANDRLSKIGKGTLHVQAKGENQGSISVGDGTVILDQQADDKGKKQAFSEIGLVSGRGTVQLNADNQFN
PDKLYFGFRGGRLDLNGHSLSFHRIQNTDEGAMIVNHNQDKESTVTITGNKDIATTGNNNSLDSKKEIAYNGWFGEKDTTKTNGRLNLVYQPAAED
RTLLLSGGTNLNGNITQTNGKLFFSGRPTPHAYNHLNDHWSQKEGIPRGEIVWDNDWINRTFKAENFQIKGGQAVVSRNVAKVKGDWHLSNHAQAV
FGVAPHQSHTICTRSDWTGLTNCVEKTITDDKVIASLTKTDISGNVDLADHAHLNLTGLATLNGNLSANGDTRYTVSHNATQNGNLSLVGNAQATF
NQATLNGNTSASGNASFNLSDHAVQNGSLTLSGNAKANVSHSALNGNVSLADKAVFHFESSRFTGQISGGKDTALHLKDSEWTLPSGTELGNLNLD
NATITLNSAYRHDAAGAQTGSATDAPRRRSRRSRRSLLSVTPPTSVESRFNTLTVNGKLNGQGTFRFMSELFGYRSDKLKLAESSEGTYTLAVNNT
GNEPASLEQLTVVEGKDNKPLSENLNFTLQNEHVDAGAWRYQLIRKDGEFRLHNPVKEQELSDKLGKAEAKKQAEKDNAQSLDALIAAGRDAVEKT
ESVAEPARQAGGENVGIMQAEEEKKRVQADKDTALAKQREAETRPATTAFPRARRARRDLPQLQPQPQPQPQRDLISRYA

SEQ ID 37 – App domain derivative

NTLTVNGKLNGQGTFRFMSELFGYRSDKLKLAESSEGTYTLAVNNTGNEPASLEQLTVVEGKDNKPLSENLNFTLQNEHVDAGAWRYQLIRKDGEF
RLHNPVKEQELSDKLGKAEAKKQAEKDNAQSLDALIAAGRDAVEKTESVAEPARQAGGENVGIMQAEEEKKRVQADKDTALAKQREAETRPATTAF
PRARRARRDLPQLQPQPQPQPQRDLISRYA

SEQ ID 38 – App domain derivative

NSGLSEFSATLNSVFAVQDELDRVFAEDRRNAVWTSGIRDTKHYRSQDFRAYRQQTDLRQIGMQKNLGSGRVGILFSHNRTENTFDDGIGNSARLA
HGAVFGQYGIDRFYIGISAGAGFSSGSLSDGIGGKIRRRVLHYGIQARYRAGFGGFGIEPHIGATRYFVQKADYRYENVNIATPGLAFNRYRAGIK
ADYSFKPAQHISITPYLSLSYTDAASGKVRTRVNTAVLAQDFGKTRSAEWGVNAEIKGFTLSLHAAAAKGPQLEAQHSAGIKLGYRW

SEQ ID 39 – App domain derivative

NTLTVNGKLNGQGTFRFMSELFGYRSDKLKLAESSEGTYTLAVNNTGNEPASLEQLTVVEGKDNKPLSENLNFTLQNEHVDAGAWRYQLIRKDGEF
RLHNPVKEQELSDKLGKAEAKKQAEKDNAQSLDALIAAGRDAVEKTESVAEPARQAGGENVGIMQAEEEKKRVQADKDTALAKQREAETRPATTAF
PRARRARRDLPQLQPQPQPQPQRDLISRYANSGLSEFSATLNSVFAVQDELDRVFAEDRRNAVWTSGIRDTKHYRSQDFRAYRQQTDLRQIGMQKN
LGSGRVGILFSHNRTENTFDDGIGNSARLAHGAVFGQYGIDRFYIGISAGAGFSSGSLSDGIGGKIRRRVLHYGIQARYRAGFGGFGIEPHIGATR
YFVQKADYRYENVNIATPGLAFNRYRAGIKADYSFKPAQHISITPYLSLSYTDAASGKVRTRVNTAVLAQDFGKTRSAEWGVNAEIKGFTLSLHAA
AAKGPQLEAQHSAGIKLGYRW

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
6 February 2003 (06.02.2003)

PCT

(10) International Publication Number
WO 03/010194 A3

(51) International Patent Classification[7]: C07K 14/22, A61K 39/095

(21) International Application Number: PCT/IB02/03396

(22) International Filing Date: 26 July 2002 (26.07.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

0118401.9	27 July 2001 (27.07.2001)	GB
0121591.2	6 September 2001 (06.09.2001)	GB
0211025.2	14 May 2002 (14.05.2002)	GB

(71) Applicant *(for all designated States except US)*: **CHIRON SRL.** [IT/IT]; Via Fiorentina, 1, I-53100 Siena (IT).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **ARICO, Maria** [IT/IT]; Chiron SpA, Via Fiorentina, 1, I-53100 Siena (IT). **COMANDUCCI, Maurizio** [IT/IT]; Chiron SpA, Via Fiorentina, 1, I-53100 Siena (IT).

(74) Agents: **MARSHALL, Cameron, John** et al.; Carpmaels & Ransford, 43 Bloomsbury Square, London WC1A 2RA (GB).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
2 October 2003

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: MENINGOCOCCUS ADHESINS NADA, APP AND ORF 40

(57) Abstract: NadA, App and ORF40 function as adhesins in *N.meningitidis*. Adhesion can be modulated by targeting these three proteins. NadA allelic variants are disclosed. Autoproteolytic cleavage of App is disclosed, as is removal of the activity by mutagenesis. App is processed and secreted into culture medium when expressed in *E.coli*. Mature App proteins are disclosed. Knockout mutants are disclosed. Vesicles from non-Neisserial hosts with heterologous adhesin expression are disclosed.

WO 03/010194 A3

PCT/IB 02/03396

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C07K14/22 A61K39/095

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, SEQUENCE SEARCH, WPI Data, BIOSIS, EMBL

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	TETTELIN H ET AL: "COMPLETE GENOME SEQUENCE OF NEISSERIA MENINGITIDIS SEROGROUP B STRAIN MC58" SCIENCE, AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE,, US, vol. 287, 2000, pages 1809-1815, XP000914963 ISSN: 0036-8075	1-14
X	-& DATABASE EMBL 'Online! 15 March 2000 (2000-03-15) TETTELIN ET AL.: retrieved from EBI Database accession no. AE002548 XP002231040 cited in the application abstract	1-14

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

'A' document defining the general state of the art which is not considered to be of particular relevance

'E' earlier document but published on or after the international filing date

'L' document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

'O' document referring to an oral disclosure, use, exhibition or other means

'P' document published prior to the international filing date but later than the priority date claimed

'T' later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

'X' document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

'Y' document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

'&' document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
13 February 2003	17.06.03

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Stolz, B

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL 'Online! 11 November 1999 (1999-11-11) FRASER ET AL.: retrieved from EBI Database accession no. AAY75736 XP002231038 abstract ---	1-14
A	WO 01 52885 A (PIZZA MARIAGRAZIA ;RAPPUOLI RINO (IT); CHIRON SPA (IT); GIULIANI M) 26 July 2001 (2001-07-26) cited in the application claims 1,14 ---	
P,X	WO 01 64922 A (COMANDUCCI MAURIZIO ;PIZZA MARIAGRAZIA (IT); CHIRON SPA (IT); GALE) 7 September 2001 (2001-09-07) cited in the application examples 9,22 -----	2-4,6-14

INTERNATIONAL SEARCH REPORT

tional application No.
PCT/IB 02/03396

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 30,33,37
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-14

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

Invention 1

Allelic variant of NadA as defined claims 1 to 14 as far as they relate to Seq ID 1

Inventions 2-14

Allelic variants of NadA as defined by claims 1 to 14 as far as theyrelate to Seq IDs 2-14.

Invention 15

claims 15 to 17 relating to methods of purifying App

Invention 16

claims 18 to 21, 25 relating to methods for preventing attachement of Neisserial cells.

Invention 17

claims 22 and 23 relating to fragments of App, NadA and ORF40

6. Claim : invention 18

claim 29 relating to a screening method

7. Claim : invention 19

claim 31 relating to a composition comprising E. coli

8. Claim : invention 20

claims 32 to 37, relating to methods for the preparation of non-Neisseria OMVs

Invention 21

claims 26 to 28, 38 to 44, and 45-51 relating to various mutants of App, NadA and ORF40.

This latter group of inventions comprises many different mutants each one of which will constitute an individual contribution.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 30,33,37

Present claim 30 relates to products defined by reference to a desirable characteristic or property, namely their being identified by the method of claim 29.
The claims cover all products having this characteristic or property, whereas the application provides no support within the meaning of Article 6 PCT and disclosure within the meaning of Article 5 PCT. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the product by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, the search has been carried out for those parts of the claims which appear to be clear, supported and disclosed, namely those parts relating to the products/compounds/methods/apparatus

Present claims 33 and 37 relate to an extremely large number of possible compounds. Support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT cannot be provided simply by reference to further pending patent applications. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

Information on patent family members

In ...al Application No
PCT/IB 02/03396

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0152885 A	26-07-2001	AU 2875401 A	31-07-2001
		CA 2397508 A1	26-07-2001
		EP 1248647 A1	16-10-2002
		WO 0152885 A1	26-07-2001
WO 0164922 A	07-09-2001	AU 3249200 A	17-11-2000
		AU 3947801 A	12-09-2001
		AU 3948801 A	12-09-2001
		CA 2371032 A1	09-11-2000
		CA 2400562 A1	07-09-2001
		CA 2400570 A1	07-09-2001
		CN 1359426 T	17-07-2002
		EP 1185691 A1	13-03-2002
		EP 1261723 A2	04-12-2002
		EP 1259627 A2	27-11-2002
		WO 0164920 A2	07-09-2001
		WO 0164922 A2	07-09-2001
		WO 0066791 A1	09-11-2000